The brain is the heaviest limb in the body. Inspire the brain to stretch the body.

BKS Iyengar (1918-)

A pair of substantial mammary glands have the advantage over the two hemispheres of the most learned professor's brain in the art of compounding a nutritive fluid for infants.

Chief Justice Oliver Wendell Holmes (1809-1894)

It's almost gone, the night is dissolving, in a cup God lifts to toast the lightning; Lightly tapping, it's high-pitched and it hums; ... There's so much less to this than you think.

Jeff Tweedy (1967-)

University of Alberta

QM and QM/MD Simulations of the *Vinca* Alkaloids

by

Evan Bradley Thomas Kelly

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science in Chemical Physics

Department of Chemistry

Edmonton, Alberta
Spring 2008

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-45830-3
Our file Notre référence
ISBN: 978-0-494-45830-3

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

The *Vinca* alkaloids are a class of pharmaceutically relevant binary-indole-indoline alkaloids based on and including natural extracts of the periwinkle plant, *Catharanthus rosea*. Two natural products, Vinblastine and Vincristine, have been in use as important chemotherapy agents for over four decades. Two semi-synthetic *Vinca* alkaloids, Vindesine and Vinorelbine, are currently in use in chemotherapy programs, and a third semi-synthetic, Vinflunine, is in advanced clinical trials. In addition to these five compounds named, which form the topic of this work, there are hundreds of other natural and semi-synthetic *Vinca* alkaloids known, although most are not clinically useful. The *Vinca* alkaloids are anti-mitotic agents and cause apoptosis instead of mitosis. It is well established that the *Vinca* alkaloids affect the cellular protein tubulin and bind to a specific site known as the *Vinca* domain. While the *Vinca* domain is well established, the specific binding mode of each drug is not. However, there is much insight into the binding mode and this has provided a strong base of information to begin simulations and to make comparisons against. Complicating the issue, however, is the large size of the *Vinca* alkaloids and their complex molecular structure, including a rotatable single bond joining the indole and indoline portions of each compound. The differential geometric and tubulin-binding properties of the drugs are not fully known. At the semi-empirical AM1 level, the projection of the potential energy surface on the major torsional angle, mentioned above, was calculated through *in vacuo* geometry optimizations. QM/MD simulations were performed, with the drugs at the AM1 level, of each *Vinca* alkaloid free in TIP3P water, and also bound to beta tubulin. A single equilibrium structure, resembling a known crystallographic Vinblastine structure [1], for the free drugs was found. Further, the 1Z2B crystal structure [2] of Vinblastine bound to tubulin appears to be a valid starting point for simulations of all five *Vinca* alkaloids studied here. This study may prove useful in future rational design of new *Vinca* alkaloids.

To

My mom, Betty; and my dad, Ross – my first science teacher

Acknowledgements

While at the University of Alberta, I spent the majority of my time working within the group of Mariusz Klobukowski. I must first of all thank Mariusz for being an excellent supervisor who made the research process enjoyable again. Mariusz is a rare example of a science professor who has balance in his life and I am sincerely proud to have been his student and to now call him my friend. To the members of the Klobukowski group over the past few years: John Lo, Jonathan Mane, Melissa Gajewski, Toby Zeng, and Jesse Kadosh: I was very happy to have you as my classmates, colleagues, and friends. I am saddened to leave our little 'dysfunctional family'. I wish you all the best.

My co-supervisor, Jack Tuszynski, is a professor who understands what the world of science will become, and he knows how to accomplish admirably big things. I thank him for the lessons he taught me in science, business, and life. To the Tuszynski group members Torin Huzil and Tyler Lutchko I offer gratitude for the many helpful discussions we had.

Most importantly, I must thank my family: Ross, Betty, Ryan, Patrick and Jessica, and of course, Gummy. This thesis is only one example of the many accomplishments that you have helped me with. Please know that I am eternally grateful for you all: without you I would not have gotten very far in this life.

Table of Contents

1 Introduction **1**

1.1 Discovery and History of the *Vinca* Alkaloids 1

1.1.1 Double Discovery 1

1.1.2 Isolation of Other *Vinca* Alkaloids 3

1.1.3 Semi-synthetic *Vinca* Alkaloids 3

1.2 Chemical Profile of the *Vinca* Alkaloids 4

1.2.1 Chemical Structures 4

1.2.2 Structural Data 8

1.2.3 Acid-Base Chemistry of the *Vinca* Alkaloids 11

1.3 Tubulin and Microtubules 12

1.3.1 Properties of Microtubules 12

1.3.2 Role of GTP and GDP 14

1.3.3 Microtubules as a Chemotherapeutic Target 15

1.4 Action of the *Vinca* Alkaloids on Microtubules 16

1.4.1 The *Vinca* Domain 16

1.4.2 Differential Binding Properties of Vinflunine 22

1.4.3 Tubulin Isotypes 23

1.5 Pharmacology of the *Vinca* Alkaloids 24

1.5.1 Current Clinical Use of *Vinca* Alkaloid Drugs 26

1.5.2 Dosages and Toxicities 26

1.5.3 Metabolism 26

1.5.4 Chemical Reactions Between VLB and Tubulin 27

1.5.5 Application to Rational Drug Design 28

2 Theoretical Introduction **29**

2.1 Overview of Computational Methods 29

2.2 Molecular Mechanics 30

2.2.1 Force Fields 30

2.2.2 TIP3P Water Model 32

2.3 Semi-empirical Quantum Mechanical Methods 33

2.3.1 Hartree-Fock Theory 33

2.3.2 CNDO 35

2.3.3 Subsequent Models 37

2.3.4 AM1 . . . 38
2.3.5 Recent Models . . . 38
2.3.6 Evaluation of the AM1 Method . . . 39
2.4 Miscellaneous . . . 40
2.4.1 Charge Fitting Methods . . . 40
2.4.2 The QM/MM Hybrid Method . . . 43
2.4.3 Energy Minimization and Molecular Dynamics . . . 45
2.5 The Blocking Method . . . 46

3 Free *Vinca* Alkaloids **48**
3.1 Torsional Potential Energy Surfaces About the Main Dihedral Angle . . . 49
3.1.1 Procedure . . . 49
3.1.2 Results . . . 50
3.2 QM/MD simulations . . . 60
3.2.1 Procedure . . . 63
3.2.2 Results . . . 66

4 *Vinca* Alkaloid-Tubulin Complexes **81**
4.1 QM/MD simulations . . . 81
4.1.1 Procedure . . . 81
4.1.2 Results . . . 82
4.2 Functional Group Residue Interactions . . . 83
4.2.1 Procedure . . . 87
4.2.2 Results . . . 87

5 Discussion **90**
5.1 Free Form *Vinca* alkaloids . . . 90
5.1.1 Potential Energy Surface . . . 90
5.1.2 MD simulations . . . 91
5.1.3 Conformation Populations . . . 92
5.1.4 *in vacuo* Dipole Moments . . . 93
5.2 The *Vinca* Alkaloids bound to Beta Tubulin . . . 93
5.2.1 MD Simulations . . . 93
5.2.2 Drug-protein pair-wise interactions . . . 94
5.2.3 Comparison of Bound to Free Form . . . 96
5.2.4 Validity of Model . . . 97
5.3 Conclusions . . . 97
5.4 Future Work . . . 98

Bibliography **99**

List of Abbreviations

ADMET	*Absorption, Distribution, Metabolism, Excretion, Toxicity model*
AM1	*Austin Model 1*
bcc	*Bond Charge Correction*
CNDO	*Complete Neglect of Differential Overlap*
ESP	*Electrostatic Potential method*
GAFF	*General Amber Force Field*
GDP	*Guanosine Diphosphate*
GTP	*Guanosine Triphosphate*
MD	*Molecular Dynamics*
MNDO	*Modified Neglect of Differential Overlap*
PDB	*Protein Data Bank*
QM	*Quantum Mechanics*
QM/MD	*Hybrid Quantum Mechanical Molecular Mechanical Dynamics*
RESP	*Restrained Electrostatic Potential*
RFO	*Rational Function Optimization*
SCF	*Self Consistent Field*
VCR	*Vincristine*
VDE	*Vindesine*
VFL	*Vinflunine*
VLB	*Vinblastine*
VNO	*Vinorelbine*

Chapter 1

Introduction

1.1 Discovery and History of the *Vinca* Alkaloids

The *Vinca* alkaloids are a very important component of many modern chemotherapy treatments. They have a unique history beginning with their near-concurrent discovery by two separate research groups, to being among the first ever tested chemotherapy drugs, to spawning the development of a whole series of successful semi-synthetic congeners. The next sections will first present the story of how the *Vinca* alkaloids were discovered and then step through the major points of their development into a series of important chemotherapeutic agents.

1.1.1 Double Discovery

The following historical introduction of the discovery of the *Vinca* alkaloids is a summation of the account by Duffin [3]. For more details of this interesting tale, the reader is referred to that article.

In 1947, at the University of Western Ontario (UWO) in London, Ontario, in a small research lab, Robert Noble, a medical doctor, and Charles Beer, an organic chemist, began a search for compounds that could be used as anti-diabetes drugs. Upon a suggestion from a personal friend, Dr. C. D. Johnston, of Jamaica, they began investigating the Madagascar red periwinkle plant as a source of active chemicals. The periwinkle plant, then known as *Vinca rosea*, is now known as *Catharanthus rosea*. The periwinkle plant was suggested because of its use in traditional Jamaican medicine. In fact, the periwinkle is used in the traditional medicine of many other parts of the world with a variety of applications. It is now known that the periwinkle plant is rich in alkaloids,

which are, broadly speaking, naturally occurring organic amine compounds. When the UWO researchers injected tea made from periwinkle leaves into test rats, they were met with unusual results: the rats died due to massive decreases in white blood cell counts and significant bone marrow depression. Noble and Beer realized the potential application of a chemical that would cause such effects as an anti-leukemia drug and began the process of isolating the active chemical in the periwinkle leaf. Through a painstaking and long process of first isolating and then testing one compound after another the team finally isolated the active compound, which was initially called Vincaleukoblastine (VLB). The name of the compound would later be shorted to vinblastine. They began performing initial anti-cancer tests, but these were limited as the group was small and without many resources. In 1955 and 1957, the first results of unpurified periwinkle extracts were published. Beer developed a method for purification of VLB and the first publication of purified VLB was in 1958. Interestingly enough, the first presentation of the VLB paper was at a cancer symposium and present in the audience were members from a research team from the large pharmaceutical company Eli Lilly. The team at Eli Lilly was involved in a massive screening project, where natural compounds were tested for anti-cancer activity at a rate of approximately 5000 every year. Included in the screening program, were extracts from the periwinkle plant – they were also included based on the use of periwinkle plants in traditional medicine, but the Eli Lilly team received their inspiration from the traditional medicine of the Phillipines. By the time of the UWO group's presentation in 1958, the Eli Lilly team had separately discovered the anti-tumour effects of periwinkle extracts, but had not published their results and also had not successfully isolated VLB. They had, however, isolated a compound similar to VLB, called Leurosine, but Leurosine was not a very potent drug and development was stopped at an early stage. Eli Lilly had resources that far exceeded the small Canadian group from UWO, but the UWO team had struck first and had the rights to the discovery of VLB and to the VLB purification method. Eventually, the two groups began to loosely cooperate, with the Eli Lilly team allowed the details of Beer's purification method, and the UWO team receiving a supply of VLB from Eli Lilly for further research.

By 1960 several clinical trials were underway. Despite the excitement felt for the promise of the *Vinca* alkaloids, chemotherapy was largely an untested method at the time, with radiation therapy being the standard cancer treatment, and thus there was a barrier of skepticism that needed to be overcome and this slowed initial testing. In fact, VLB was one of the first cancer drugs to undergo clinical trials. By this point, it had been discovered that the mechanism of action of VLB was that of a 'spindle poison',

where the drug would cause cell death during mitosis by interfering with the mitotic spindle apparatus. In 1962, Eli Lilly released to the public information about their discovery of a new *Vinca* alkaloid from the periwinkle plant named vincristine (VCR). Vincristine was originally thought to be more potent than vinblastine, but also appeared to be accompanied by neurotoxic effects not observed with vinblastine.

While the clinical trials of VLB and VCR administered alone were impressive and demonstrated the ability of *Vinca* alkaloids to block the proliferation of cancer cells [4], the first grand successes of the *Vinca* alkaloids and chemotherapy in general came in 1965 with the development of the combination chemotherapy POMP[1] method, where VCR was concurrently administered in a 'cocktail' with methotrexate, 6-mercaptopurine (6-MP) and prednisone. The POMP regimen had great success in forcing remission in childhood leukemia, and this was one of the first effective cures for cancer [5].

1.1.2 Isolation of Other *Vinca* Alkaloids

With the successes of VLB and VCR proven in leukemia, there is great interest in the discovery of other *Vinca* alkaloids that would be effective against leukemia or other types of cancer. At least 130 compounds have been isolated from the periwinkle plant, and 40 of them are similar in structure (see Figs. 1.1 and 1.2) to VLB and VCR. There has been great effort expended over the past several decades to test other compounds isolated from the periwinkle plant. However, currently only two natural *Vinca* alkaloids are approved for clinical use: VLB and VCR. An interesting example is the naturally occurring compound Vincovaline (VCV), which, differing only in its stereochemistry at two sites, is an epimer of VLB. VCV has shown almost none of the anti-mitotic activity of VLB, despite its similarity to VLB [6]. A reduced form of VLB, 3'-4'-Anhydrovinblastine, is a naturally occurring *Vinca* alkaloid that is reported to be in phase II clinical trials as of 2005 [7].

1.1.3 Semi-synthetic *Vinca* Alkaloids

With the success of VLB and VCR as chemotherapeutic agents, there has been large interest in synthetic pathways to VLB and VCR as isolation from the plant source is very inefficient. Furthermore, these synthetic routes could lead to producing similar compounds not found in nature. By virtue of its discovery being the first, VLB is considered

[1] the name 'POMP' is an acronym based on the brand names of each of the four drugs: Purinethol (6-MP), Oncovin (Vincristine), Methotrexate, Prednisone

the 'lead compound' of the *Vinca* alkaloids and congeneric compounds are discussed and named in relation to VLB. By designing and/or synthesizing VLB congeners, the mechanism of action of the *Vinca* alkaloids can be better understood and the ultimate goal is to produce compounds that are better drugs in that they have fewer harmful side effects and are more effective against cancer. Over the past decades, hundreds of semi-synthetic *Vinca* alkaloids have reportedly been produced [8]; the first semi-synthetic to reach clinical trials was vinglycinate (VLB with a glycine adduct) in 1967 [9]. Vinglycinate, like many subsequent *Vinca* alkaloid drug candidates, was not pursued because it showed no benefits over use of VLB or VCR. Many of the efforts in semi-synthetic *Vinca* alkaloid design have shown great creativity, such as vinfosiltine [10], a VLB congener with a novel aminophosphonate adduct. While vinfosiltine showed incredible *in vitro* potency, it was abandoned after initial clinical tests showed no superiority over VLB [11, 12]. Another example of a failed derivative is vinepidine [11, 13], which proved too toxic for clinical use: toxicity is another common problem with *Vinca* alkaloids. Unfortunately, the documentation regarding the synthesis, structure, and *in vitro* and *in vivo* testing of semi-synthetic *Vinca* alkaloids is not always complete or publicly available.

The first success of a semi-synthetic *Vinca* alkaloid was deacetyl vinblastine amide, or vindesine (VDE), Fig. 1.3, which was produced at Eli Lilly in 1978 using VLB as a starting material [14]. Vindesine is currently only approved for clinical use in Europe. The second semi-synthetic to be approved for clinical use was 5'-noranhydrovinblastine, or vinorelbine (VNO). It was designed by the group of Pierre Potier after they devised a synthetic scheme for the total synthesis *Vinca* alkaloids [15, 16]. Vinorelbine was first approved for clinical use in France in the late 1980s and by the mid-1990s in the USA. The 3rd semi-synthetic drug to be considered in this thesis is a vinorelbine derivative named vinflunine (VFL), which is currently in phase III clinical trials as of 2004. Vinflunine was synthesized using novel superacid flourination techniques [11].

1.2 Chemical Profile of the *Vinca* Alkaloids

1.2.1 Chemical Structures

The chemical structures of the *Vinca* alkaloids studied in this thesis are presented here. Fig. 1.1 shows the structure of VLB. All other drugs will be discussed in relation to VLB. Differences compared to VLB are shown in either red or green: red denotes a change or conversion, and green denotes a site of a deletion. The structure of these *Vinca* alkaloids is that of a binary indole-indoline compound: the indole or 'top' portion of the drug is

Table 1.1: Nomenclature of *Vinca* alkaloids used in this thesis

Common Name	Trade Name	Abbrevation
Vinblastine	Velban	VLB
Vincristine	Oncovin	VCR
Vindesine	Eldisine	VDE
Vinorelbine	Navelbine	VNO
Vinflunine	Javlor	VFL

known as cleavamine or carbomethoxyvelbenamine, and the indoline or 'bottom' portion is known as vindoline. Rai [17] cautions that the name 'catharanthine' for the top portion is not correct. As most derivitizations are usually restricted to either the top or the bottom portion of the molecules it can be useful to separate derivatives into two categories: 1) vinorelbine and vinflunine are top derivatives; 2) vincristine and vindesine are bottom derivatives. Fig. 1.2 shows VCR, which differs in the conversion of a methyl group to a formyl group. Vindesine is shown in Fig. 1.3 and involves the conversion of two ester groups to an amide and a hydroxyl. Vinorelbine, Fig. 1.4, involves a reduction across a C-C bond and also the contraction of the 9-membered ring of VLB to an 8-membered ring. Vinflunine, Fig. 1.5 involves the same ring contraction as vinorelbine and also the addition of two flourine atoms.

Several atom naming conventions for the backbone atoms of the binary *Vinca* alkaloids have been used in the scientific literature. Here, the IUPAC convention, which is the most common, will be adopted. The atom names are shown for VLB in Fig. 1.6. The naming convention is easily transferred between the five VLB congeners under study here. The only exception is in VNO and VFL where the nine-membered ring in the top portion of the molecules is shrunk: in this case, the C8' atom will be considered to have been deleted.

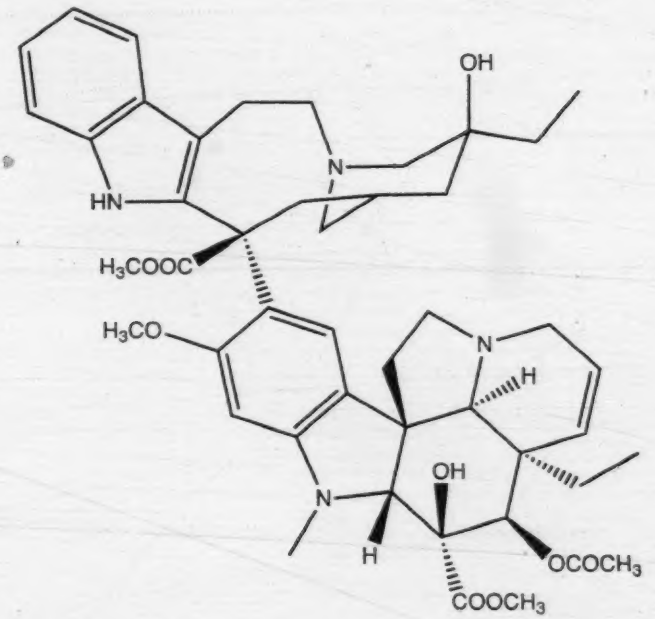

Figure 1.1: The chemical structure of vinblastine.

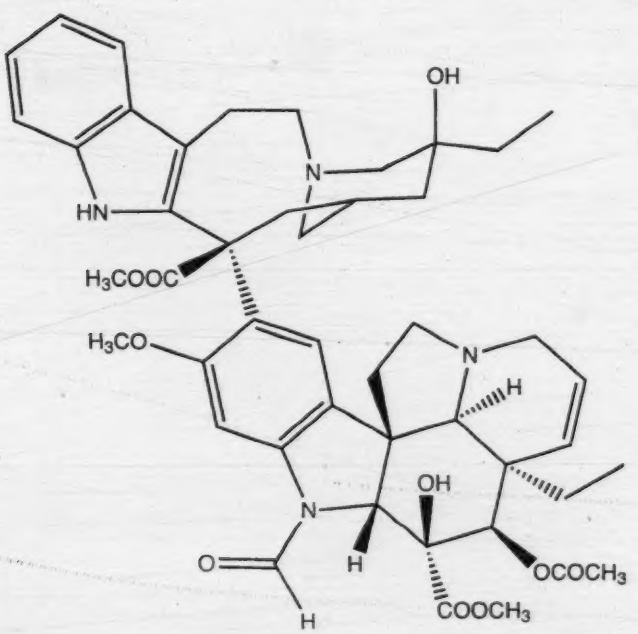

Figure 1.2: The chemical structure of vincristine.

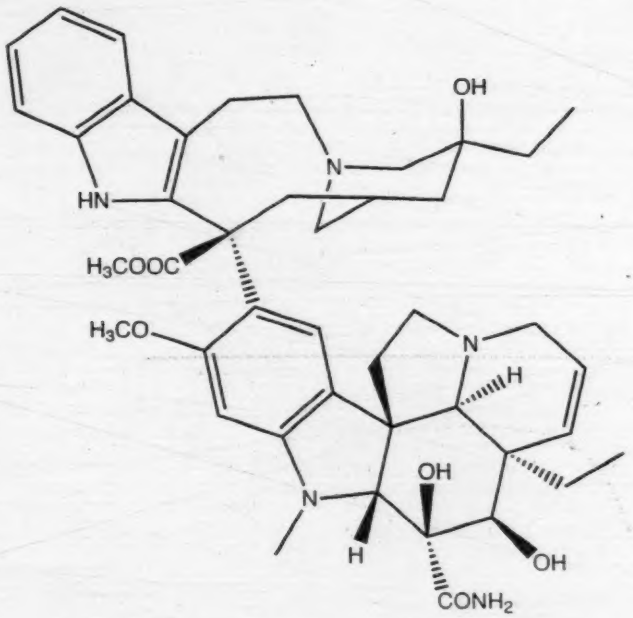

Figure 1.3: The chemical structure of vindesine.

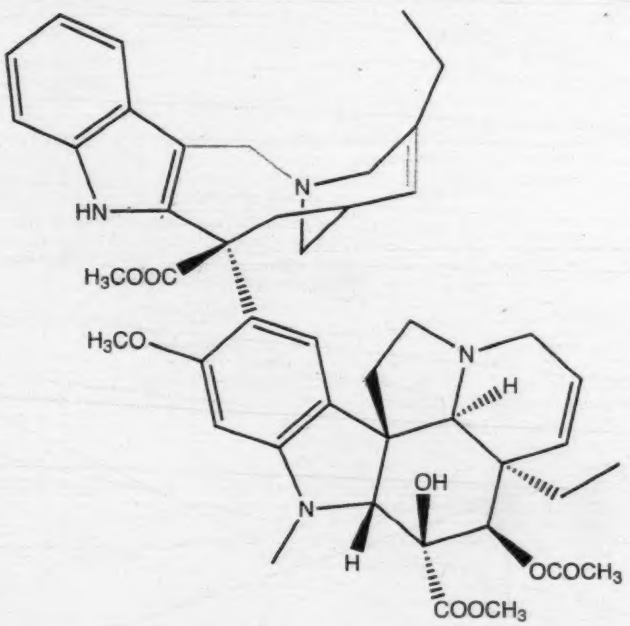

Figure 1.4: The chemical structure of vinorelbine.

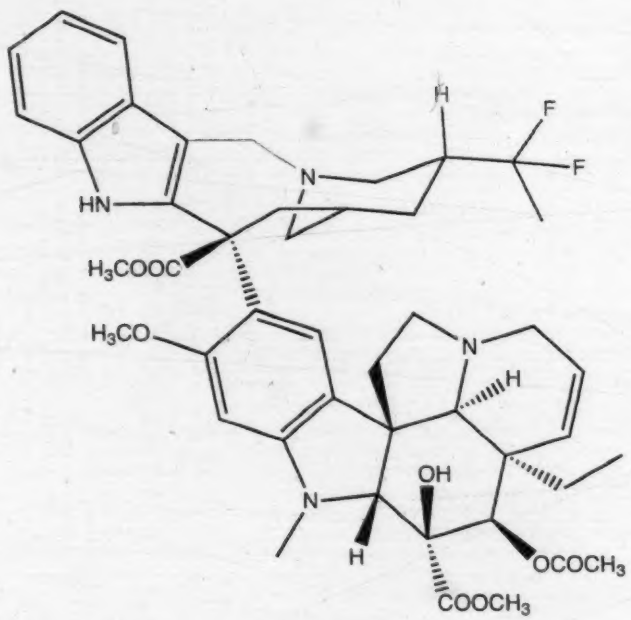

Figure 1.5: The chemical structure of vinflunine.

1.2.2 Structural Data

There have been a variety of structural and geometric studies performed on the *Vinca* alkaloids. This section will present several studies, in chronological order, and summarize the most important and relevant points pertaining to molecular geometry and structure. Of particular interest is the C17'-C18'-C15-C16 dihedral angle (referred to as the 'main dihedral angle', see Fig. 1.6) about the single C-C bond joining the top and bottom portions of the drug.

In 1965, Moncrief and Lipscomb obtained an X-ray crystal structure of vincristine methiodide [18]. They found the main dihedral to be approximately 160°. A similar crystal structure for VLB was not found: the crystal structure of VLB has proven challenging to discover.

In 1983, 2D-NMR studies of VLB separately in both benzene and dichloromethane at room temperature were presented by Hunter [19]. A hydrogen bond between N9 and the C3 hydroxyl proton was observed. The main dihedral was estimated to be in the range of 140-170°. No explanation is given of how the structural analysis and modeling

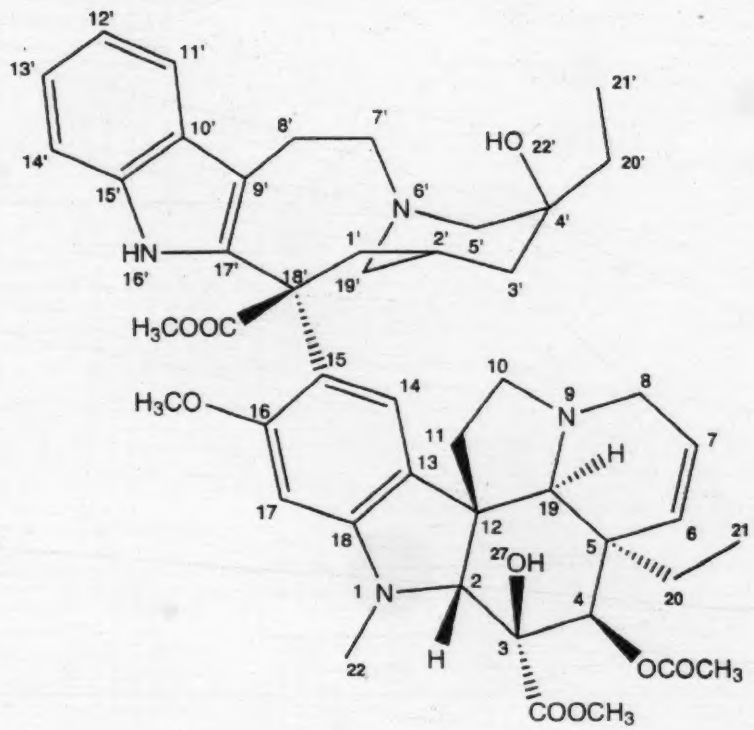

Figure 1.6: The atom naming system for vinblastine.

were done, and it is assumed that no computational models were constructed – perhaps a physical 'ball and stick model' reflecting the distance constraints found was the basis of their geometric analysis.

In 1992, Gaggelli [20] performed a 2D-NMR proton and ^{13}C NMR study of VLB in aqueous solvent. The pH of the solution was between 4.6 and 6.6. The temperature was not stated explicitly and therefore assumed to be room temperature. A computational model was constructed at the MM2 [21] force field level and an energy-minimized structure was found under NMR distance restraints. The main dihedral angle was found to be approximately 180°.

In 1995, Andrews [22] presented a 2D-NMR study of 3'-4'-anhydrovinblastine in trichloromethane at 263K. Only peak assignments were made from the NMR data and distance restraints were not discussed. The MacroModel 4 [23] program was used to perform a Monte-Carlo search for minimum geometry structures at the MM3* [24–26] force field level. The lowest energy structure found was one with a main dihedral angle of 37° and the second lowest energy structure had an angle of 207°. These structures were discussed in terms of the NMR data and the 207° structure was considered to be in the closest agreement with experiment.

In 2000, Bau and Jin [1] applied new crystallization techniques to VLB and found an X-ray crystal structure of vinblastine sulfate at 173K. The geometry of this structure was very similar to Moncrief's structure of VCR from 35 years earlier. A hydrogen bond between N9 (which is protonated in this structure as the crystal was formed from acidic solution) and O27 was identified – this bond is not relevant if the molecule is not in acidic media. The main dihedral angle in Bau and Jin's structure was 162°. They also quote an unpublished result of a main dihedral angle of about 165° from an X-ray crystal structure of vinblastine chloride.

In 2001, Ribet [22] performed an NMR study of VFL in acetone. Only peak assignments were made, and chemical environment discussion was made, but geometric analysis was not performed. However, the purpose of the study was to verify the chemical structure of VFL, and not the geometry.

There are several structural features of the *Vinca* alkaloids that are worth mentioning. They are multi-cyclic binary organic compounds of two relatively large groups joined by

a rotatable single C-C bond. They are large molecules of approximately 800 Da, each containing approximately 60 heavy (non-hydrogen) atoms. The 8- or 9-membered rings in a bi-cycle with a 6 membered ring, found in the top portions of the molecules, are unusual. There is an abundance of stereogenic centres, adding to the complexity of the molecules. The Bau and Jin structure, and as well the Gaggelli and Moncrief structures, find that the C1'-C7'-C8'-C18' portion of VLB and VCR is flat and nearly coplanar with the flat C9'-C17' indole moiety. These studies also find the piperidine ring in the top portion is in a stable chair conformation. The structural analysis by Gaggelli is considered to be the most relevant to this study as it is done in aqueous solution in only slightly acidic conditions near pH 7.

1.2.3 Acid-Base Chemistry of the *Vinca* Alkaloids

There are four nitrogen atoms available for lone-pair donation on each of the binary *Vinca* alkaloids. The bonded neighbours of each of these weakly basic nitrogen atoms in the five *Vinca* alkaloids included in this study are similar, thus the acid-base nature of each is expected to be similar. It has been found [27] that there are two basic atoms on each of VLB and VCR, with pK_a values of 5.4 and 7.4, and 5.0 and 7.4, respectively. No data is available for the other three compounds. The crystal structure of VLB [1] shows two sites of nitrogen protonation, and is thus in agreement with the above pKa values. However, the experimental conditions in [1] were those of a crystal being grown from acidic solution. The main consideration for this study is the behaviour of the drugs at human physiological conditions, which is typically of pH slightly greater than 7. Chemical equilibrium theory gives the percent ionized of a base as:

$$percent\ ionized = \frac{100}{1 + 10^{pH - pK_a}} \tag{1.1}$$

which gives the result that under slightly basic conditions, approximately 50% of the *Vinca* alkaloid drug will be protonated at the site of pK_a 7.4. If a second protonation were to occur, then a approximately further 10% would be protonated at the site of pK_a 5.0 or 5.4. This means that only a small fraction of the compound will be at a +2 charged state at physiological pH. The majority of a sample of a *Vinca* alkaloid in aqueous solution will be either in a neutral state or a +1 charged state, each in about equal amounts. Unfortunately, with the experimental data available, it is not possible to narrow the pK_a site down to one specific nitrogen atom. Because of this, each *Vinca* alkaloid will be modeled as a neutral molecule, and this is assumed to be a valid state to study these drugs based on the above discussion of ionization fractions.

Figure 1.7: Resonance structures of an amide moeity.

Vindesine has an amide moiety present in its structure and this is potentially a fifth basic nitrogen. However, because of electron sharing through the resonance structure shown in Fig. 1.7, the amide nitrogen lone pair is not localized on the nitrogen and thus this is not a site of significant basicity.

1.3 Tubulin and Microtubules

The *Vinca* alkaloids are known to be anti-mitotic and cause apoptosis instead of cell division. The anti-mitotic action of the *Vinca* alkaloids is likely through binding to tubulin and interference with microtubule dynamics. Thus, in this section, the basic properties of microtubule biopolymers and their monomer protein tubulin will be discussed, as well as the interaction between *Vinca* alkaloids and tubulin and microtubules. The reader is referred to an excellent review of microtubules and microtubule-targeting drugs by Jordan and Wilson [28].

1.3.1 Properties of Microtubules

Microtubules are a component of the cellular cytoskeleton and they provide structure to cells. A crucial role of microtubules is to provide mechanical force as part of the mitotic spindle apparatus during cellular mitosis. Tubulin is primarily present in microtubules in two different forms: alpha tubulin and beta tubulin: these forms almost always exist in a hetero-dimer of a beta unit and alpha unit (see Fig. 1.8). Alpha and beta tubulin have different primary amino acid sequences, but they are 40% identical at the sequence level. However, these two forms of tubulin are indistinguishable beyond a resolution of 6 Å in electron crystallography [29]. The interaction is specific in that the same surface region of an alpha tubulin unit will always interact with the same region of a beta tubulin unit. Thus, all alpha beta heterodimers are essentially the same. By convention, the beta unit will be considered to be on 'top' and the alpha unit on the 'bottom'. The hetero-dimers

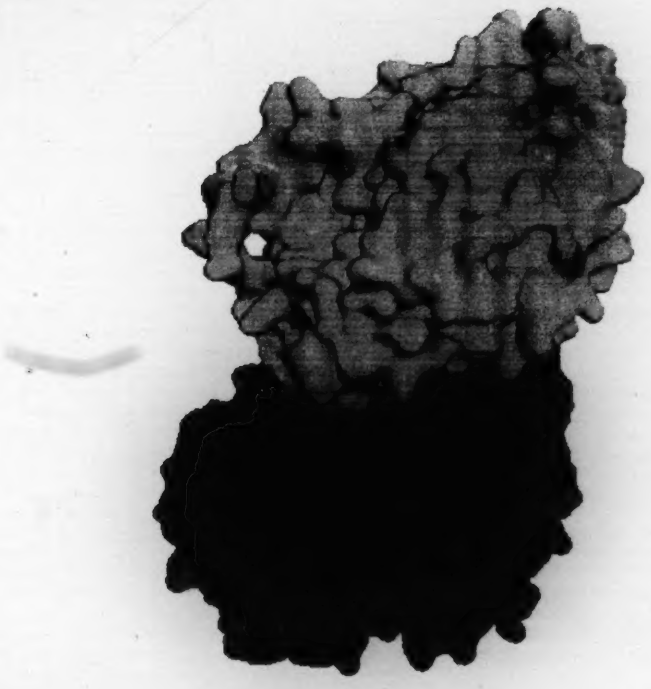

Figure 1.8: An alpha-beta tubulin heterodimer. The coordinates were taken from the 1JFF PDB structure [30] and is presented in a VdW surface rendering. The red (bottom) region is the alpha unit, and the green (top) region is the beta unit. Figure created using VMD [31].

are held together by non-covalent interactions. One dimer has a molecular weight of about 100 kDa and a size of approximately 4 nm by 5 nm by 8 nm.

Tubulin heterodimers can interact end-to-end and form long protofilaments. These protofilaments then can interact laterally with each other and can form sheets, which in turn can roll into a cylinder. The resulting long, hollow cylinder with a circumference of approximately 13 heterodimers is what is known as a microtubule.

Microtubules display complex dynamic behaviour: once a microtubule is nucleated, it initially grows rapidly in length with the addition of heterodimer subunits. The growing end is known as the '+end' of the microtubule; the other end, that was the site of nucleation, is the '-end'. The inter-dimer interaction is purely non-covalent. The +end of a microtubule is very dynamic in that it is the site of subunit addition resulting in growth

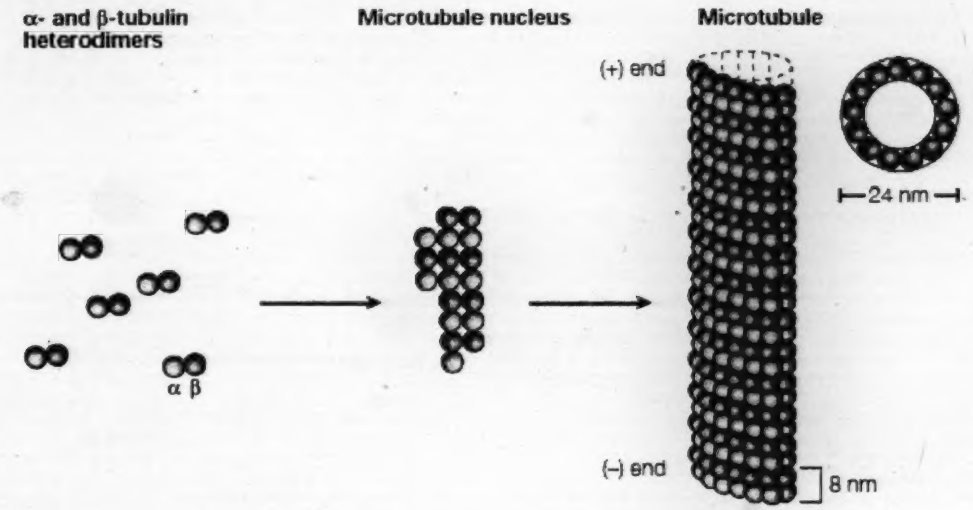

Figure 1.9: The formation of a microtubule. Taken from [28]

and subunit loss resulting in shortening. The -end, while it does undergo similar dynamic shortening and growth, is more stable than the +end and the length change rate at the -end is much smaller. The interchanging growth/shortening behaviour of microtubules is referred to as 'dynamic instability' [32, 33]. A 'catastrophe' is when a microtubule shifts from growth to rapid disassembly; a 'rescue' is the reverse change from shrinking to growth. The dynamic instability of microtubules is very important as the role of microtubules in mitosis requires microtubules to rapidly rearrange themselves in order to accommodate cellular shape changes as the cell divides. An interesting dynamic behaviour of microtubules is 'treadmilling', where one end grows and the other end shrinks, both at the same rate. Treadmilling arises from differential tubulin concentrations at the microtubule ends.

1.3.2 Role of GTP and GDP

GTP, guanosine triphosphate, can be a source of chemical energy through its hydrolysis to GDP, guanosine diphosphate, and an inorganic phosphate group (PO_4^{3-}), P_i, as shown in Fig. 1.10.

There is hydrolysable GTP present in microtubules in a stoichiometry of one GTP per tubulin heterodimer. There are other GTP or GDP molecules present, but they are

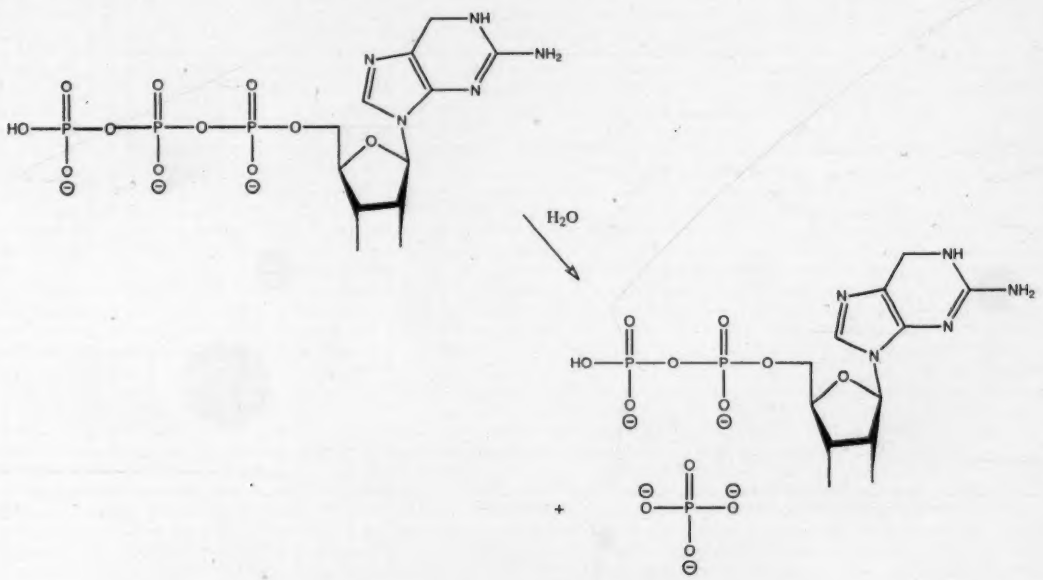

Figure 1.10: The hydrolysis of GTP.

not hydrolysable, or at least not in processes relevant to microtubule dynamics. The hydrolysable GTP site is near the tip of the beta-tubulin portion of a heterodimer, and is thus at the +end of the microtubule. The general understanding of the role of GTP in microtubule dynamics is that GTP hydrolysis will occur during or shortly after the addition of a new heterodimer to the tip of a microtubule. Thus, most hydrolysable GTP sites in the microtubule contain GDP and not GTP. If the microtubule tip contains primarily GTP or GDP-P_i complexes, it is 'capped' and in a stable conformation allowing for further growth. If hydrolysis of GTP and the loss of the P_i unit occurs, a conformational change occurs and the microtubule +end is destabilized. This loss of the GTP cap and resultant conformational change may result in catastrophe [28, 34].

1.3.3 Microtubules as a Chemotherapeutic Target

Interference, even if subtle, with microtubule dynamics can disrupt the critical role of microtubules during mitosis. A disruption will often lead to apoptosis (programmed cell death) instead of mitosis. The interference may potentially be in the form of a ligand bound to a microtubule causing the stabilization of microtubules preventing catastrophe, or in the destabilization of microtubules preventing growth, or both of these effects together. Because cancer cells generally divide faster than healthy cells, the initiation of

apoptosis instead of mitosis is a desirable goal for an anti-cancer treatment: the cancer cells, which undergo mitosis more frequently than normal cells will then die at a higher rate. Thus, if a drug indiscriminately causes cell death at mitosis, this is potentially the basis of an effective cancer treatment. However, such an anti-mitotic drug would then kill healthy cells as well as cancerous cells. Even though the healthy cells would die at a slower rate than the rapidly proliferating cancer cells, dangerous side effects would likely then be seen. It is, therefore, a goal in anti-mitotic drug choice or design to target cancerous cells more strongly than healthy cells in order to reduce side effects and provide a more effective treatment.

1.4 Action of the *Vinca* Alkaloids on Microtubules

At high (μM) concentrations of *Vinca* alkaloids in *in vitro* experiments, the action of the *Vinca* alkaloids is usually to completely depolymerize microtubules [28]. However, at lower (nanomolar and sub-nanomolar) concentrations, the *Vinca* alkaloids leave microtubules largely intact [34]. In fact, the nanomolar effect appears to be stabilization: microtubule dynamics is slowed and microtubules spend more time in a paused state with no growth or attenuation. The effect of the *Vinca* alkaloids then appears to be twofold: the net rate of microtubule polymerization is reduced and the rate of depolymerization is also reduced [35]. It is likely that it is at this low concentration that the *Vinca* alkaloids act at the cellular level *in vivo*. It is interesting to note that studies of vinblastine have shown that only one or two molecules of *Vinca* alkaloids are needed per microtubule to induce the above mentioned stabilization effect [34].

1.4.1 The *Vinca* Domain

Generally speaking, VLB is known to interact with the +end of a microtubule. But, is the site of interaction localized? The answer is yes, and several studies have determined the site. This interaction site is known as the '*Vinca* domain', and is the general site of interaction on tubulin for the *Vinca* alkaloids [28], but also for many other compounds [7] such as rhizoxin, maytansine, the cryptophycins, and the dolastatins, to name a few. Competitive inhibition is observed between the above mentioned compounds, thus confirming that they all bind to the same site on tubulin. The *Vinca* domain is a high-affinity site for the *Vinca* alkaloids to interact with and this binding is rapid and reversible [36]. Several of the most important studies on the determination of the *Vinca* domain will be presented below.

In 1996, Rai [17] performed a photochemical study of the binding of VLB to unpolymerized rat brain tubulin. They developed a derivative of VLB with a small fluorescent anthranilate adduct at the C4' position on the top portion of VLB, as shown in Fig. 1.11, which is referred to as Ant-VLB. Ant-VLB successfully inhibited microtubule polymerization, but was found to be 7.5 times less potent than VLB. Ant-VLB was found to competitively inhibit maytansine, and thus was confirmed to bind to the *Vinca* domain. Low-affinity binding sites were found at high Ant-VLB concentrations and a high-affinity site was observed for lower concentrations. The anthranilate fluorescence maximum wavelength is blue-shifted upon binding, which indicates a change in the flourophore's environmental polarity upon binding. This may be an indication that the C4' region of VLB is buried upon binding. Photolabelling experiments with Ant-VLB revealed that the high-affinity binding site is localized to beta tubulin and residues β175-β213 were identified as comprising the *Vinca* domain. The photolabeling was accomplished by irradiation at 325 nm, which should only excite the anthrilate adduct, and not tubulin, to cause covalent adduction. This study places the *Vinca* domain at the tip of a tubulin heterodimer, which coincides with the observation that the *Vinca* alkaloids bind to the tips of the +end of microtubules.

In 2002, a second photochemical study was performed by Chatterjee [37]. Here the flourophore, 3-carbonylamino-7-diethylaminocoumarin, was attached to the bottom portion of VLB at the ester group at C4. The resulting compound is referred to as F-VLB. VLB was found to inhibit binding of F-VLB to tubulin, and both had roughly the same affinity for tubulin as the K_{eq} values for binding were similar. This is confirmation that F-VLB does bind to the *Vinca* domain. The authors believe that the fluorescent adduct on F-VLB is solvent exposed when bound to tubulin, based on Stokes' shift observations. F-VLB inhibited microtubule polymerization in a similar fashion to VLB, and thus the adduct at C4 is hypothesized to not be involved in binding to tubulin, which supports the result that the adduct does not directly contact tubulin and remains solvent exposed upon binding. This is perhaps in contrast to Anti-VLB, where the fluorescent adduct is thought to be in contact with tubulin and the anti-polymerization effects decreased compared to VLB. The distance from the *Vinca* domain to the binding site for colchicine was found to be approximately 40 Å, which is consistent with the residue binding map determined by Rai.

Mitra [38], in 2004, performed a computational docking study on several anti-mitotic peptides that are all known to bind at the *Vinca* domain. Cryptophycin 1, cryptophycin

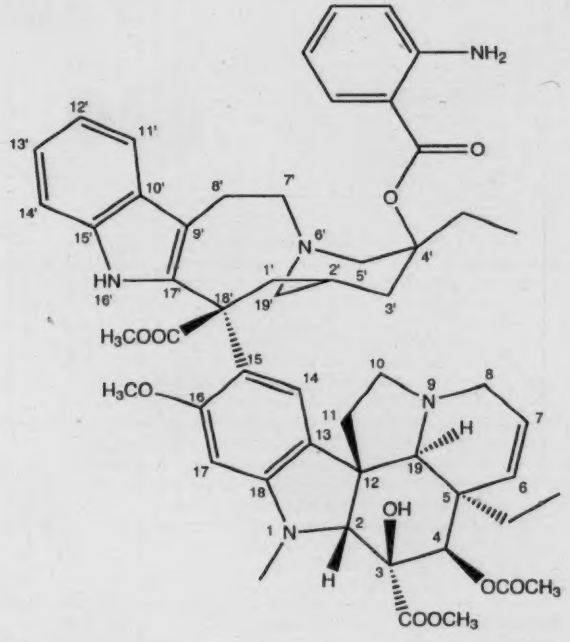

Figure 1.11: The structure of Ant-VLB, vinblastine labeled with the photoactive anthranilate moiety at the C4' position.

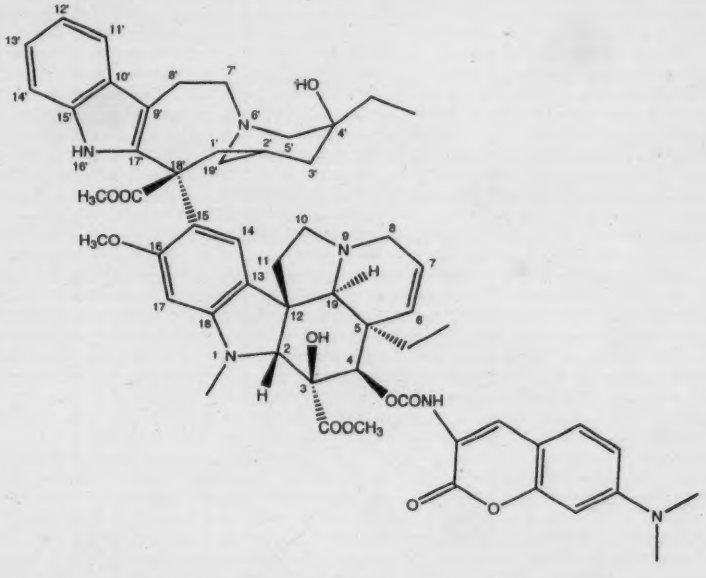

Figure 1.12: The structure of F-VLB, vinblastine labeled with the photoactive 3-carbonylamino-7-diethylaminocoumarin moiety at the C4 position.

52, dolastatin 10, hemiasterlin, and phomopsin A were all included. Docking was attempted, separately, to both alpha and beta tubulin. Beginning with structures from the 1JFF PDB [30] structure of bovine tubulin, the authors performed 2.5ns MD simulations on the tubulin units and took five coordinate 'snapshots', one every 0.5 ns, to generate an ensemble of protein configurations to use in docking. The NAMD [39] simulation program was used to perform the MD simulations with the CHARMM27 [40, 41] force field. Using AUTODOCK 3.0 [42], the peptide ligands, as described by the Tripos [43] force field and Gasteiger-Marsili [44,45] atomic charges, were docked into the ensemble of tubulin proteins using a flexible ligand and rigid protein protocol. No site of appreciable affinity was found on alpha tubulin, and one single high-affinity site was located on beta tubulin. The contact residues describing the high-affinity site were found to be: β172, β174, β175, β177, β204, β205, β208, β209, β212, β220, β222. This site is in strong agreement with the residues identified by Rai.

In 2005, Gigant [2] presented an X-ray crystal structure of a complex of two tubulin heterodimers longitudinally arranged (as in a protofilament) that were crystallized under stabilization of colchicine and with an RB3-SLD backbone. Vinblastine was then soaked into the crystal and the X-ray experiment was performed. The crystal structure was obtained to a resolution of 4.1 Å. Using C12'-bromo-VLB, an electron density difference map was constructed to unambiguously confirm the location of VLB when bound to tubulin, and the residue binding map denotes the same region as discovered by Rai [17] and by Mitra [38]. In this structure, the VLB molecule is sandwiched between the two heterodimers, contacting both the tip of one beta tubulin unit and the tip of the next alpha tubulin unit. This crystal structure, containing both bound VLB and bound colchicine is consistent with the 40 Å separation as determined by Chatterjee [37]. This study is very important as it provides atomic coordinates for VLB bound to tubulin. These coordinates may be used as a starting point for computational calculations of *Vinca* alkaloids bound to tubulin, such as molecular dynamics simulations. The main dihedral angle of VLB was found to be 213°. In the bound conformation of VLB in this structure, the C4' atom is buried in the beta tubulin unit it contacts, and the C4 atom would be solvent accessible if the second heterodimer wasn't present, as suggested by the photolabeling studies mentioned above. It is interesting to note that in this structure, the *Vinca* domain is located very close to the hydrolysable GTP site (which contains GDP in this structure): the two sites are as close as 7 Å apart (see Figs. 1.13 and 1.14). The coordinates from this study are available as the PDB structure 1Z2B.

Figure 1.13: The 1Z2B structure of a tubulin heterodimer showing VLB (red) in the *Vinca* domain and a GDP molecule(orange) in the hydrolysable GTP site. Figure created using VMD [31].

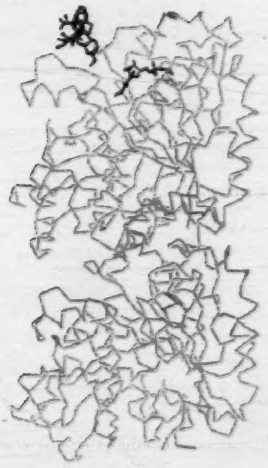

Figure 1.14: An alternate presentation of the *Vinca* domain and hydrolysable GTP site. Figure created using VMD [31].

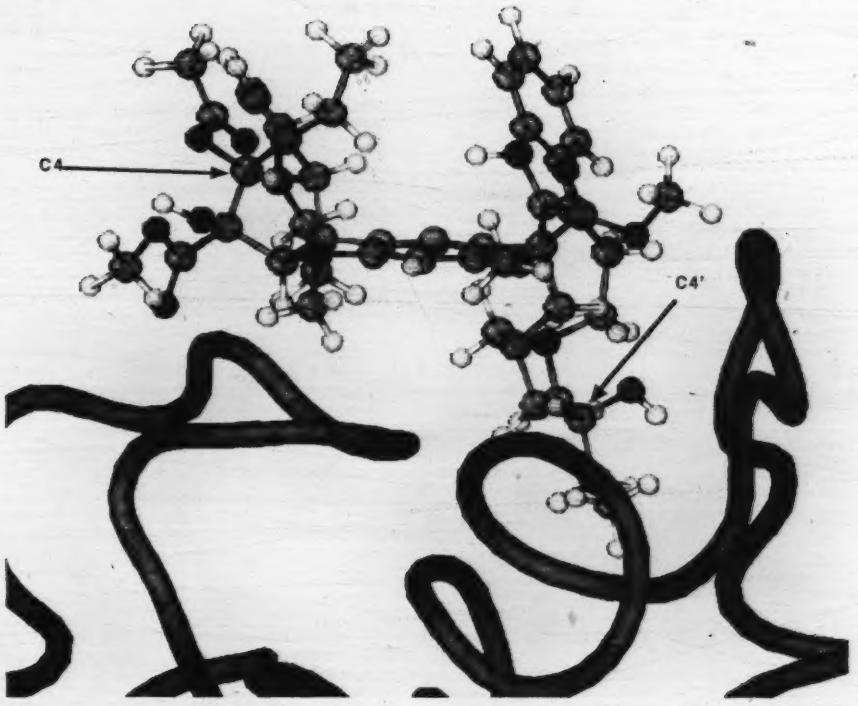

Figure 1.15: A close-up of VLB bound to tubulin in the 1Z2B crystal structure. VLB atoms C4 and C4' are labeled to indicate that C4 is solvent exposed, as suggested by Chatterjee, and C4' is buried in the protein, as suggested by Rai. Figure created using VMD [31].

In the 1Z2B study Gigant [2], the authors used fluorescence changes [46] to measure rate constants for the binding of VLB to the stabilized tubulin complex, and from these results they proposed a two step mechanism consisting of: 1) rapid formation of a collision complex; and, 2) a slow rearrangement.

The four studies presented above demonstrate that the *Vinca* domain is well defined and identified with a reasonably high degree of accuracy. Furthermore, the *Vinca* domain is known to be only a single site and not a collection of several binding sites. This single binding domain can then be assumed with confidence in this thesis.

This single, high-affinity binding site is likely the main cause for the anti-mitotic action of *Vinca* alkaloid drugs as that site is likely the only one significantly active at lower, non-toxic concentrations. However, it is worth mentioning that at high *Vinca* alkaloid concentrations, many low-affinity sites are present throughout a microtubule. There is also weak interaction between the *Vinca* alkaloids and the -end of a microtubule [36]. However, because it is the most physiologically relevant, the focus must be on the high-affinity +end site.

1.4.2 Differential Binding Properties of Vinflunine

It is not surprising that each of the five *Vinca* alkaloids discussed in this study behave similarly in their interactions with tubulin and microtubules in solution. Competitive inhibition studies clearly demonstrate that each of the *Vinca* alkaloids bind to the *Vinca* domain [47]. However, the effect of each drug is subtly different. For example, studies show that each will bind with a different affinity for microtubules, resulting in differing IC_{50} (the concentration of drug that inhibits microtubule assembly by 50%) values for each drug. A relative ordering of binding strengths, not including VDE, was found by Kruczynski [47] to be, from most strongly bound to weakest bound: VCR > VLB > VNO > VFL. Thus, it is reasonable to assume that while the mechanism and binding of the *Vinca* alkaloids are largely the same, the minor chemical differences between them cause subtle changes in the binding of each to tubulin. However, Kruczynski also found that VFL's binding properties made it stand out from the other three *Vinca* alkaloids in her study. First, VFL does inhibit tubulin polymerization as the other *Vinca* alkaloids do, and this is the most important characteristic of a *Vinca* alkaloid as an anti-mitotic drug. Also, VFL was shown to interact with the *Vinca* domain and, as evidenced by its displacement from binding by all three of VLB, VCR, and VNO. However, in the reverse experiment, VFL would not displace the other three drugs, even at high VFL

concentrations (100 μM). Furthermore, a VFL-tubulin complex was not detectable in the study, where complexes were detectable for the other three *Vinca* alkaloids. This presents VFL in an interesting light: while it produces the same effects on microtubules as the other *Vinca* alkaloids, its binding behaviour is different, in some way, from the other *Vinca* alkaloids. Bearing in mind that the other *Vinca* alkaloids are thought to bind in a rapid and reversible manner, a *Vinca* alkaloid drug is expected to undergo a dynamic cycle of repeated complexation and dissociation. It has been hypothesized [11] that, based on Kruczynski's results, the rate of dissociation is increased for VFL. Also, the binding strength may be decreased. This would mean that VFL spends less time in a bound state. In competitive inhibition experiments, other *Vinca* alkaloids would be able to easily occupy the binding sites, left open for a majority of the time by VFL, and as their rentention times are longer, they would effectively displace the VFL molecules from binding. The cause for this binding difference is not evident, but it may involve a different kind of interaction of VFL with tubulin, such as a different binding mode at the *Vinca* domain.

1.4.3 Tubulin Isotypes

Proteins that are homologous in their primary amino acid sequence are referred to as isotypes. Isotypes are essentially the same protein and fulfill the same principal roles in the human body (or in other forms of life). Sequence deviations can be neutral or essential: neutral deviations cause no change in function and essential deviations adapt the protein to a specific function. Different isotypes of tubulin are expressed in different degrees in different human organs and cells. Recent work in the Tuszynski group has revealed ten relevant human isotypes of beta tubulin [29], and other work has shown six isotypes of alpha tubulin [28]. For the purpose of studying the *Vinca* alkaloids the isotypes of beta tubulin are most important as that is the site of the *Vinca* domain. Tubulin isotype mutations may directly affect microtubule dynamics through differential chemical or electrostatic interactions, or indirectly through conformational changes . Drug binding may be affected for the same reasons. For example, it is known that VLB binds best to the βII isotype [29]. Also, cancer cells are known to express certain isotypes of tubulin as compared to their healthy counterparts. Therefore, there exists the potential to design isotype-specific drugs, which would target the overexpressed cancer isotypes and thus act with greater potency and with less toxic side effects. As seen in the isotype sequences in Fig. 1.16 and visualized in Fig. 1.17, there are very few isotype mutations in the *Vinca* domain, but isotype specificity of the *Vinca* alkaloids may be accomplished due to indirect effects of sequence differences elsewhere in the protein

```
Residue:    172.....180   208.....225
bovine      SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaI       SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaIIa     SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaIIb     SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaIII     SPKVSDTVV.....YDICFRTLKLATPTYGDL
betaIVa     SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaIVb     SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaV       SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaVI      SPKVSDTVV.....YDICFRTLKLTTPTYGDL
betaVII     LPKVSDTVV.....YDICSRTLKLPTPTYGDL
betaVIII    SPKVSDTVV.....YDICSKTLKLPTPTYGDL

RED amino acids are divergent mutations
BLUE amino acids are conserved mutations
```

Figure 1.16: Amino Acid Variation in the *Vinca* Domain as identified by Huzil [29]. Divergent and conservative mutations are labeled as such strictly on the basis of amino acid side-chain comparison and do not necessarily correlate with neutral and essential isotype designations.

due to allosteric effects. The scope of this study does not permit the study of all ten human beta isotypes, thus the bovine brain isotype is chosen as the sequence for beta tubulin used. The bovine sequence is identical to seven of the human isotypes in the *Vinca* domain and bovine brain tubulin is used for the majority of experimental studies done.

1.5 Pharmacology of the *Vinca* Alkaloids

Drug binding assays in aqueous solution, as discussed above, provide useful information about the action of a drug. However, the true test of a drug is its effectiveness when administered to a human. The human body is a complex organism and binding to a target protein is the last of many steps for a drug, thus drug effectiveness is not necessarily easy to predict. A standard model for a drug in the human body is the ADMET model:

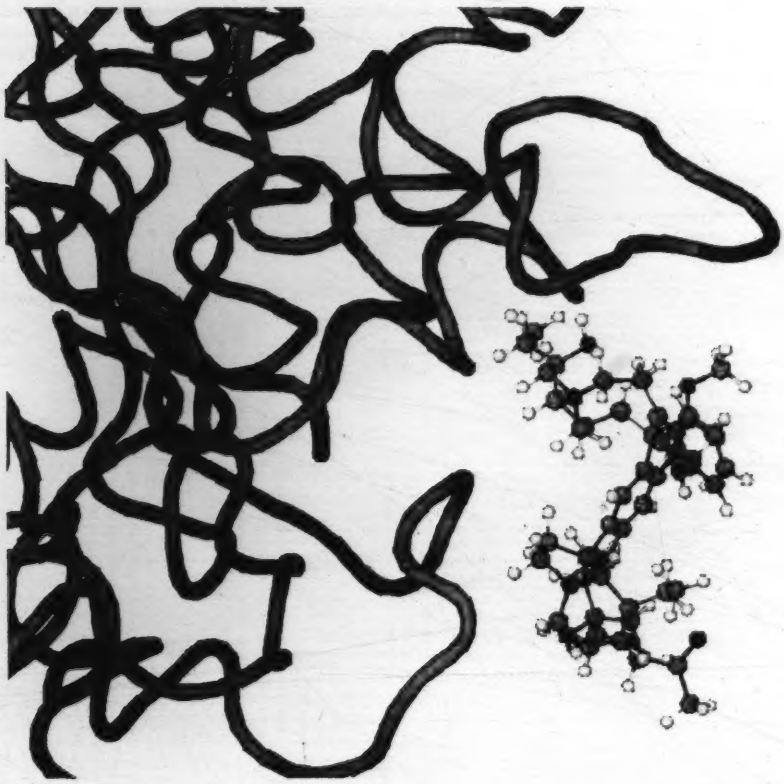

Figure 1.17: VLB in the *Vinca* domain, from a snapshot in the QM/MD simulation trajectory. Sites of isotype mutations are shown in red. Figure created using VMD [31].

Absorption, Distribution, Metabolism, Excretion, and Toxicity. Each of these aspects plays an important role in the effectiveness of a drug and must be considered. The following sections will briefly discuss the pharmacology of the *Vinca* alkaloids.

1.5.1 Current Clinical Use of *Vinca* Alkaloid Drugs

Vinblastine is typically used in cases of Hodgkin's disease and testicular germ-cell cancer. Vincristine is used in cases of leukemia and lymphomas. Interestingly, childhood tumors are more sensitive to the effects of VCR than to VLB, thus VCR is a common drug in pediatric oncology [48]. Vindesine has found a use as part of a cocktail against non-small-cell lung cancer [11]. Vinorelbine has been successful against solid tumours, lymphomas, and lung cancer. Vinflunine has shown promising results in clinical trials against cancer of the bladder, non-small-cell lung cancer and breast cancer [28]. There are numerous combination cocktails in use involving *Vinca* alkaloids and there are many more being tested in clinical trials.

1.5.2 Dosages and Toxicities

Along with their differing clinical spectra, the administration and toxic side-effects of each *Vinca* alkaloid are different. Some pertinent pharmacological data are presented in Table 1.2. The neurotoxicity of VCR is likely due to the interaction of the drug with neuronal microtubules – microtubules are prevalent throughout the body, which allows many opportunities for unwanted side-effects of *Vinca* alkaloid drugs. It is worth mentioning that VLB and VNO have been observed to induce neurotoxic effects, but the severity of that particular side-effect with these two drugs is much less than with VCR. The more efficient clearance of VNO from the body is considered a reason to allow VNO to be administered in higher doses than the other *Vinca* alkaloid drugs, which can allow for a more effective treatment.

1.5.3 Metabolism

Each of the four clinically active drugs is know to undergo hepatic metabolism to some degree. Metabolites could be simple hydrolysis products, or products of enzymatic cleavage. Metabolites can be difficult to experimentally identify and quantitate and as result the process of metabolism is not well understood and there are potentially undiscovered metabolites for each *Vinca* alkaloid. 4-Desacetyl-VLB is a known major metabolite of VLB [49], and studies have suggested that this metabolite is up to five times more potent as a microtubule inhibitor than VLB [50]. Desacetyl-VNO and N-oxide-VNO are both

Table 1.2: Pharmacological information for the *Vinca* alkaloids [48]

	VLB	VCR	VDE	VNO
Adult Dosage[a]	6-8	1-2	3-4	15-30
Clearance[b]	0.74	0.16	0.25	0.40-1.29
Dose Limiting Factor	Neutropenia[c]	Neurotoxicity	Neutropenia	Neutropenia

[a]value in $mg/m^2/week$
[b]value in L/kg/h
[c]Neutropenia is a reduction in certain types of white blood cells

identified as potential metabolites of VNO, but are contrasted in that Desacetyl-VNO has roughly the same level of *in vivo* activity as its parent compound, and N-oxide-VNO is essentially inactive [51]. Elmarakaby discussed a very unique enzymatic metabolite [52], called Catharinine. In Catharinine, the C4'-C5' bond of VLB is cleaved, resulting in the C4' hydroxyl group being oxidized to a carbonyl group, and the C5' atom being oxidized to a formyl group. The piperidine ring of VLB is then opened in Catharinine. It was found that Catharinine is 77 times less effective than VLB in *in vitro* tests.

With the knowledge that both the parent drug and potentially several metabolites will be present in the body upon administration of a *Vinca* alkaloid drug, a very important question may be posed: which compounds are active? It may be the case that the parent compounds are entirely inactive *in vivo* and a metabolite causes the actual anti-mitotic effects. It could also be the case the parent compound doesn't cause any toxic side-effects, but rather a metabolite does. To fully understand the activity of the *Vinca* alkaloids, it is then crucial to understand their individual metabolic fates and the activities of their metabolites.

1.5.4 Chemical Reactions Between VLB and Tubulin

It is interesting to note a proposal made by Magnus [53], that VLB may act as a methylating agent of thiol groups – in particular, the -SH groups of cysteine residues on tubulin. There are several chemotherapeutic agents, including Cucurbitacin B, Taxodione, and Brefeldin, that act as alkylating agents on enzymes or other proteins; alkylation can be

a source of interference of enzyme or protein function [54]. The methylation reaction is initiated by reversible protonation at the ipso position of the benzene ring of the bottom portion of VLB (i.e. protonation of the C15 atom). This leads to the cleavage of the C15-C18' bond in VLB and the activation of the N1 methyl group. It is this methyl group that will then be added to a thiol group. While VLB is able to undergo this reaction, the conversion of the N1 methyl group to a formyl group in VCR disables this reaction and thus VCR cannot act as a methylating agent. VCR is reported [55] to be metabolized less in the body than VLB is, which may be a result of the differing availability of this specific reaction. This idea is interesting and relevant since it raises the question of anti-mitotic action of the *Vinca* alkaloids: is it based in a chemical reaction or does it arise from non-covalent ligand-protein binding, or some combination of both?

1.5.5 Application to Rational Drug Design

It is important to note that these drugs are all chemically very similar and the differences noted in clinical uses, administration, metabolism, and side-effects are all due to minor changes from one drug to the next. With this idea of small modifications causing noticeable changes on the drug's properties, extra motivation is gained for the study and design of new semi-synthetic congeners of VLB.

Chapter 2

Theoretical Introduction

2.1 Overview of Computational Methods

The potential energy, U, of a chemical system is of fundamental importance to theoretical and computational methods. One simple reason for this is that the classical force on an object can be expressed as the gradient of the potential energy with respect to the object's spatial coordinates:

$$\mathbf{F} = -\nabla U. \tag{2.1}$$

Objects tend to move in the direction of the net force on them towards a minimum point of their potential energy – the force on an object is zero when at a potential energy minimum. In the same fashion, a chemical system will tend to follow the various forces on its components towards a minimum of its potential energy surface (the hypersurface of potential energy with respect to the spatial coordinates of the system). With knowledge of the potential energy surface and/or forces of a system, it is then possible to perform energy optimizations, molecular dynamics simulations, and Monte Carlo simulations to calculate various properties of the system, such as molecular geometries and dipole moments. Some of these methods will be discussed in more detail later in this chapter, but suffice it to say that the method of finding the potential energy of a chemical system is important and is worth discussion.

There are two main methods of calculating the potential energy of a chemical system. One is the classical treatment of molecular mechanics (MM) where electrons are not explicitly modeled. Molecular mechanics has the benefit of being reasonably simple and computationally efficient, but since electrons are not explicitly included it lacks the

general ability to model electronic effects, such as bond formation and breaking. If electrons are included explicitly in the model, the theory of quantum mechanics (QM) must be the basis of the method due to the small mass of an electron. Quantum mechanical models give a generally more accurate description of a system and allow for electronic effects such as bond formation and breaking, polarization, and the study of excited states. However, quantum mechanical models are more complicated than molecular mechanical models and require much more computational power and time. The modeling of even modestly-sized systems comprised of a few hundred atoms can be of prohibitive computational expense. Both quantum mechanical and molecular mechanical methods have their place in the computational chemistry field and each has a place in the research at hand. The remainder of this chapter will discuss the relevant methods and some of the important details of each involved.

2.2 Molecular Mechanics

The following is a summary of molecular mechanics theory and methods and is not intended to be a full reference. For a more complete account, the reader is referred to a computational chemistry book such as [56].

2.2.1 Force Fields

Within molecular mechanics, the potential energy function is the foundation of the method, but this function can be varied in its formation and many different functions exist. These functions are often referred to as 'force fields' as the forces on a system are easily found from the potential energy functions used. To begin with, most methods treat each atom as a point charge and point mass. The potential energy function is usually assumed to take the form of a sum of various separate components:

$$U = U_{bond} + U_{bend} + U_{torsion} + U_{elec} + U_{VdW} \tag{2.2}$$

where U_{bond} is a term to describe stretching of chemical bonds, U_{bend} for bending of the angle between two consecutive bonds, $U_{torsion}$ is to describe torsion of a dihedral angle, U_{elec} is the electrostatic interaction between pairs of atoms, and U_{VdW} describes the Van der Waals interactions between pairs of atoms. To find the form of each of these terms, U_{bond} is used as an example: its form is found by taking a Taylor series of U with respect to the stretching of a bond length r, about r_o, the equilibrium bond length:

$$U(r) = U(r_o) + U'(r_o)(r - r_o) + \frac{1}{2}U''(r_o)(r - r_o)^2 + \frac{1}{6}U'''(r_o)(r - r_o)^3 + \dots \quad (2.3)$$

where the equilibrium energy, $U(r_o)$, can be arbitrarily set to zero and the first derivative of U at the equilibrium bond length is also zero since, by assumption, r_o is a minimum of U with respect to r. Thus, the first two terms of (2.3) vanish. Then, the choice of which order of Taylor polynomial is used to describe bond stretching must be made. Most often, only the quadratic term is kept and that describes the potential energy of a chemical bond near its equilibrium point very well. However, cubic and quartic terms are sometimes kept. A cubic term will introduce anharmonicity into the bond potential energy well, which reflects the asymmetry of bond stretching farther away from the equilibrium point: as two atoms become close the potential energy curve is steep, but when the atoms stretch far apart the curve flattens out. The standard bond stretching term, for an entire system, keeping only the quadratic term is:

$$U_{bond} = \frac{1}{2}\sum_{bonds} k_{AB}(r_{AB} - r_{AB,o})^2 \quad (2.4)$$

where the sum is over all pairs, A and B, of chemically bonded atoms present, k_{AB} is the force constant for each bond, and $r_{AB,o}$ is the equilibrium bond length for each bond. The constants k_{AB} and $r_{AB,o}$ depend solely on the types of the atoms A and B bonded together. A polynomial form such (2.4) is desirable because it is simple, continuous everywhere, differentiable to all orders, and all derivatives are continuous. A similar Taylor series analysis about a bond angle, θ, will lead to:

$$U_{bend} = \frac{1}{2}\sum_{angles} k_{ABC}(\theta_{ABC} - \theta_{ABC,o})^2 \quad (2.5)$$

where the sum is over all bond angles formed by a triple of consecutive bonded atoms, A, B and C. As with the bonding case, the k_{ABC} and $\theta_{ABC,o}$ values are force constants and equilibrium bond angles that depend on the atoms types of A, B, and C. For the torsion term in (2.2) a Fourier series is used instead of a Taylor series due to the periodic nature of torsional rotation. Using a cosine Fourier series and keeping only the first two terms the standard term is:

$$U_{bend} = \frac{1}{2}\sum_{dihedrals} V_{ABCD}[1 + (-1)^{n+1}cos(n\omega_{ABCD} + \phi_{ABCD})] \quad (2.6)$$

where the sum is over all dihedral angles formed by a quadruple of consecutive bonded

atoms, A, B, C, and D. V_{ABCD} is the amplitude of the energy curve for the atom quadruple $ABCD$, n is the periodicity of the torsion, ω_{ABCD} is the dihedral angle and ϕ_{ABCD} a phase angle. The electrostatic interaction is accomplished by assigning partial charges, q_A, to each atom and using Coulomb's law:

$$U_{elec} = \sum_{pairs\ A,B} \frac{q_A q_B}{r_{AB}} \tag{2.7}$$

where the sum is over all pairs of non-bonded atoms A and B. While the Van der Waals term may take several different functional forms, the most common is the Lennard-Jones '12-6' function:

$$U_{VdW} = \sum_{pairs\ A,B} \left(\frac{a_{AB}}{r_{AB}^{12}} - \frac{b_{AB}}{r_{AB}^{6}}\right) \tag{2.8}$$

where the sum is over all non-bonded pairs and a_{AB} and b_{AB} are constants depending on the atoms types of A and B. Equation (2.2) thus defines a general force field and the definitions of its terms given above describe a simple field that is suitable for standard systems at or near equilibrium where no bonding changes will occur. More complicated terms may be used to create a more robust field or for more exotic applications.

While the above equations set out the mathematical framework for a force field, the question of how each of the many parameters involved above are determined is still left unanswered. An entire set of parameters must be defined before a force field can be of any use. Parametrizations can either be intended for use in general systems or can be tuned for a specific system. Caution must be exercised in choosing a force field as its parametrization may be entirely unsuitable for the system under study and the results of a calculation or simulation will be unreliable. Most parametrizations rely on statistical fitting to either experimental data or *ab initio* calculations or a combination of both.

2.2.2 TIP3P Water Model

Often, it is desirable to model a system as a solute in some solvent as many properties are solvent specific and would not be modeled properly *in vacuo*. For biologically relevant molecules, the solvent is most commonly water. One method of including aqueous solvent effects in a model is to explicitly include water molecules surrounding the solute. It is easily imagined that hundreds or thousands of individual water molecules would be required to adequately surround a solute. This demands that a water model should not only be simple and efficient, but also robust and accurate. Thus, a model of a water

molecule must be carefully constructed. In cases where water molecules will not be involved in any chemical bonding with the system, a force-field representation is the clear choice under consideration of efficiency. One of the simplest and most obvious models is to represent a water molecule as three connected point charges: two hydrogen atoms and one oxygen atom. Each water molecule must be neutral and this is accomplished by placing an identical positive charge on each hydrogen atom and a negative charge on the oxygen atom equal to twice the charge magnitude on one hydrogen atom. It is worth mentioning that it is possible to add one or more off-atom charges to the model in an effort to better represent the charge distribution of the water molecule. A simplification to this force field model is then to constrain the bond lengths and the bond angle in the water molecule. The only interaction parameters then needed are those for the intermolecular electrostatic and dispersion forces. One of the most successful and widely used 3-point water models, as described above, is the TIP3P model [57]. In TIP3P, the OH bond length is 0.9572 Å, the HOH bond angle is 104.52°, the charge on each hydrogen atom is $0.417e$, and the charge on the oxygen atom is $-0.834e$. These values were optimized in order to reproduce the density of liquid water, as well as the energetics and geometry of gas-phase complexes and of liquid water. Four and five centre water models of a similar style, TIP4P and TIP5P, respectively, are also available [57] and are typically more accurate than TIP3P but are not as often used since they are more complicated and slower. The TIP3P model is generally considered to be over-polarized [58] and this needs to be considered when applying the model.

2.3 Semi-empirical Quantum Mechanical Methods

When it becomes necessary or desirable to explicitly include the electrons of molecules in a model, the foundation of the method must be quantum mechanical and forming a simple classical potential energy function such as (2.2) is not possible. While a number of different versions of quantum mechanical methods exist to study chemistry problems, the next sections will focus on an introduction to semi-empirical methods, which are suitable for efficient study of larger systems, such as biologically relevant ligands or pharmaceutical compounds. The following is only a summary of the theory and methods and a more thorough review of quantum mechanical methods can be found in [59].

2.3.1 Hartree-Fock Theory

In order to discuss semi-empirical quantum chemical methods the basis of these methods, Hartree-Fock theory, must first be presented. This theory begins with the *ansatz* that

the multi-electron molecular wavefunction, Ψ, is represented as a Slater determinant:

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \dots & \psi_N(1) \\ \psi_1(2) & \psi_2(2) & \dots & \psi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(N) & \psi_2(N) & \dots & \psi_N(N) \end{vmatrix} \tag{2.9}$$

where N is the number of electrons in the system and each $\psi_i(j)$ is the i^{th} spin-orbital of the j^{th} electron's coordinates. A spin orbital is the product of a spatial orbital, ϕ, and a spin-eigenfunction, α (for spin up) or β (for spin down). An example of a spin-orbital is:

$$\psi_i(j) = \phi_i(j)\alpha(j) \tag{2.10}$$

where each spatial orbital is expanded in terms of a basis set comprised of functions χ_μ:

$$\phi_i(j) = \sum_\mu C_{i\mu}\chi_\mu(j). \tag{2.11}$$

To obtain the restricted closed-shell Hartree-Fock theory, each orbital is filled by two electrons with opposite spins. The electronic structure of a molecule is determined from the matrix Hartree-Fock-Roothaan-Hall equation:

$$\mathbf{FC} = \mathbf{SCE} \tag{2.12}$$

where $\mathbf{F}$ is the Fock matrix, with elements:

$$\begin{aligned} \mathbf{F}_{\mu\nu} = &\int \chi_\mu \left(-\frac{1}{2}\nabla^2\right)\chi_\nu \mathrm{d}r - \sum_k^{nuclei} Z_k \int \chi_\mu \left(\frac{1}{r_k}\right)\chi_\nu \mathrm{d}r \\ &+ \sum_{\lambda,\kappa} P_{\lambda\kappa}\left[\int \chi_\mu(1)\chi_\nu(1)\frac{1}{r_{12}}\chi_\lambda(2)\chi_\kappa(2)\mathrm{d}r_1\mathrm{d}r_2\right. \\ &\left. - \frac{1}{2}\int \chi_\mu(1)\chi_\lambda(1)\frac{1}{r_{12}}\chi_\nu(2)\chi_\kappa(2)\mathrm{d}r_1\mathrm{d}r_2\right]. \end{aligned} \tag{2.13}$$

In equation (2.13), $P_{\lambda\sigma}$ is an element of the density matrix P:

$$P_{\lambda\sigma} = 2\sum_i^{occupied} C_{\lambda i}C_{\sigma i} \tag{2.14}$$

where the C_{ij} are the elements of the coefficient matrix, **C**, which describes the contribution of each basis function to each molecular orbital as in equation (2.11). In equation (2.12) **S** is the overlap matrix describing the degree of coincidence of two basis functions:

$$S_{\mu\nu} = \int \chi_\mu \chi_\nu \mathrm{d}r \tag{2.15}$$

and **E** is a diagonal matrix of molecular orbital energies. If, for a particular basis set, the set of molecular orbital coefficients in **C** can be found, then the electronic structure of the system in question has been determined. However, the elements of the Fock matrix, **F**, must first be calculated and these elements depend on the very solution to the problem, the **C** matrix (through the density matrix). Thus, it is necessary to begin with a guess for the **C** matrix, from which values of **E** are found, which then allows for a new **C** matrix to be formed. This process is then continued iteratively until the results converge to one stable answer: this is called the Self-Consistent-Field (SCF) method.

2.3.2 CNDO

The Hartree-Fock SCF method described above has proven to be effective in modeling many types of chemical systems. It has the advantage over force field methods of being able to model excited electronic states, bond formation and breaking, and many other properties that require explicit inclusion of electrons. However, the computational power and time required to solve the Hartree-Fock equations is often prohibitive. If larger systems are to be studied using quantum mechanical methods, approximations must be made to the above theory in order to make the equations more efficient. One method is to replace certain terms with simpler functions relying on parameters fit to experimental data. This gives rise to the semi-empirical methods. One of the first semi-empirical methods developed that was generally applicable was the Complete Neglect of Differential Overlap (CNDO) method [60, 61]. More modern semi-empirical methods may be seen as based on CNDO and it will be used as an example to illustrate the theme of semi-empirical methods.

In CNDO the basic framework of the Hartree-Fock theory, the Hartree-Fock-Roothaan-Hall equation is kept. The Fock matrix is of special interest as it contains many terms that are time-consuming to compute and there is room for simplification. The first approximation made is the removal of all core electrons from explicit treatment; to compensate for this loss of electrons, nuclear charges are reduced by the number of core

electrons removed. The major assumption in CNDO is to not include molecular overlap integrals from different basis functions. This makes the overlap matrix defined in (2.15) the identity matrix and the Hartree-Fock-Roothaan-Hall equation (2.12) becomes a set of regular eigenvalue equations:

$$\mathbf{FC} = \mathbf{CE}. \tag{2.16}$$

Next, there are numerous simplifications made to the elements of the Fock matrix. The last two terms in (2.13), the two electron integrals:

$$\int \chi_\mu(1)\chi_\lambda(1)\frac{1}{r_{12}}\chi_\nu(2)\chi_\kappa(2)\mathrm{d}r_1\mathrm{d}r_2 \tag{2.17}$$

are set to zero if they come from different basis functions. That is, two-centre integrals are included only if $\mu = \lambda$ and $\nu = \kappa$. All remaining two-centre integrals are not calculated explicitly, but each is set to some function, γ, which is a function of only the two atom types, A and B, involved and the distance, r, between the two atoms. Although there are many possibilities for γ, one option proposed by Mataga and Nishimoto [62] is:

$$\gamma(A,B,r) = \frac{\gamma(A,A)+\gamma(B,B)}{2+r+[\gamma(A,A)+\gamma(B,B)]}. \tag{2.18}$$

For a single atom $\gamma(A,A)$ was previously defined by Pariser and Parr [63,64] as

$$\gamma(A,A) = IE(A) - EA(A) \tag{2.19}$$

where $IE(A)$ stands for ionization energy and $EA(A)$ is electron affinity for atom type A. This use of ionization energies and electron affinities is a prime example of the use of experimentally determined values to simplify Hartree-Fock theory. The first two terms in (2.13) are collectively referred to as the resonance integral, and these are approximated for integrals arising from the same basis function as:

$$\int \chi_\mu \left(-\frac{1}{2}\nabla^2\right)\chi_\mu \mathrm{d}r - \sum_k^{nuclei} Z_k \int \chi_\mu \left(\frac{1}{r_k}\right)\chi_\mu \mathrm{d}r = V(A,B). \tag{2.20}$$

In (2.20) V is a function depending on the atom types, nuclear charges involved, and potentially some experimental values such as the ionization energy. For cases where the two basis functions are different, the approximation is:

$$\int \chi_\mu(-\frac{1}{2}\nabla^2)\chi_\nu \mathrm{d}r - \sum_{k=nuclei} Z_k \int \chi_\mu(\frac{1}{r_k})\chi_\nu \mathrm{d}r = \beta_{AB}S_{\mu\nu} \tag{2.21}$$

where β_{AB} is a parameter depending only on the two atom types, and $S_{\mu\nu}$ is the overlap integral. Note that two different sets of overlap integrals are used concurrently in this method, the full set used in (2.21) and the set used in forming the Hartree-Fock-Roothaan-Hall equation where all overlap integrals arising from different basis functions are set to zero resulting in **S** becoming the identity matrix.

With all of these approximations in hand, the Hartree-Fock-Roothaan-Hall equation is then formed and solved using the same iterative procedure as in full Hartree-Fock theory. While some of the approximations used above may seem unphysical and obscure in nature, the fundamental point of CNDO to be conveyed here is that the exact terms of the Hartree-Fock-Roothaan-Hall equation are replaced by simpler functions containing parameters that are either equated to experimental observables or statistically fit to a set of experimental values such as heats of formations or molecular geometries. All semi-empirical methods use the same philosophy to provide a simplification to the Hartree-Fock theory.

CNDO was a success when compared to previous attempts at approximations to Hartree-Fock theory, such as the Hückel method [65] and the Pariser-Parr-Pople (PPP) [66–68] method in that it was able to accurately predict 3D molecular geometries and polarizations; the previous methods were restricted to planar molecules. CNDO was also considerably faster than full Hartree-Fock calculations done on computer. Unfortunately, errors were quickly found: one serious flaw being that neutral atoms were attracted to each other over large distances. Proper convergence of the iterative SCF solving technique was, also, at times hard to obtain, and that was also a major problem.

2.3.3 Subsequent Models

In an attempt to address the shortcomings in the CNDO method a variety of revisions were made. Each kept the same framework and semi-empirical concept of fitting to experimental data, but utilized more realistic functional forms, neglected fewer terms, removed restrictions of equality of integrals, and employed better statistical fitting procedures over larger sets of experimental data. These methods include, but are not limited to, in chronological order of development: NDDO (*Neglect of Diatomic Differential Overlap*) [60], CNDO/2 [69], INDO (*Intermediate Neglect of Differential Overlap*) [70], MINDO/3 [71], MNDO (*Modified Neglect of Diatomic Overlap)* [72]. MNDO, developed by Dewar in 1977, was an elegant method in the style of NDDO and proved to be useful in the calculation of many properties. It was parametrized to reproduce heats of forma-

tion, dipole moments, ionization energies, and molecular geometries. It also made the important jump in developing a flexible model and used only monatomic data in fitting resonance integral terms instead of the previously used diatomic data. MNDO was eventually parametrized to include many inorganic elements as well as the standard organic elements. However, it was found to be unable to correctly model dispersion forces, and in particular was unable to correctly model hydrogen bonds.

2.3.4 AM1

By 1985, Dewar and co-workers had developed a new version of the MNDO method, entitled *Austin Model 1*, or AM1 [73]. The improvements over MNDO included the inclusion of a core-core interaction term in the hamiltonian that fixed MNDO's inability to accurately model dispersion forces. Also, more sophisticated and powerful parametrization methods were used with AM1; the data sets used saw the beginning in a shift from using atomic data to molecular data. AM1 has become a widely used semi-empirical method. In the work described in this thesis, the AM1 method is used in two different software packages: PC-GAMESS [74] and AMBER [75]. Both packages use the original set of parameters for all atoms used in this work [73, 76].

2.3.5 Recent Models

In 1989, Stewart [77,78] presented a new NDDO-style method. Since it was third in line after MNDO and AM1, it was called *Parametric Model 3* (PM3). Stewart used almost the exact same theoretical form as AM1, but with more advanced parametrization methods and larger data sets. Continuing the shift seen in AM1, more molecular data were used and atomic data were limited. In fact, PM3 is thought of as having used the most molecular data possible for a method of its type. The basic idea behind PM3, using an AM1-like framework but with a different parametrization technique and/or larger data sets, set the stage for the majority of the work done on semi-empirical methods over the roughly two decades since its release. There have been several recent models, which although not used here, are worth brief mention. The *Recife Model 1* (RM1) [79], uses exactly the same framework as AM1, but has a new parametrization based on 1736 organic and biologically relevant molecules. The RM1 model was designed with modeling biologically relevant molecules (namely pharmaceuticals) as a goal. While it has not seen as much testing and usage as AM1, initial results suggest its performance is better in some applications and worse in others when compared to AM1 and PM3. A strong point of RM1 is that it only differs from AM1 in the parameters used and thus any of

the large number of programs that have AM1 implemented can easily be modified to use RM1. The PDDG/PM3 and PDDG/MNDO [80] methods are variants of the PM3 and MNDO methods: they include a new *Pairwise Distance Directed Gaussian* (PDDG) core-core interaction term to replace the similar AM1 solution to the problem of incorrect dispersion forces. The two PDDG methods use a special parametrization and have shown encouraging results and may have promise for use in modeling of pharmaceutical compounds due to increased subtlety in recognition of functional group differences allowed by the new PDDG. Included in Fujitsu's MOPAC7 program [81], is the PM6 method. PM6 is another reparametrization of the NDDO framework, but again with a new form of core-core interaction corrective terms. It has been parametrized against over 9000 compounds. PM6 is proprietary to Fujitsu.

2.3.6 Evaluation of the AM1 Method

The AM1 and PM3 methods are the longest lived and most used of the methods mentioned above: they are both considered to be generally reliable methods. Of particular interest here is the ability of the model to predict molecular geometry. Several results of the accuracy of geometric values and other important measures of reliability are quoted in Table 2.1 below, for organic molecules. Clearly, both AM1 and PM3 could both be chosen for most calculations or simulations involving organic molecules. AM1 was chosen for the current study, rather than PM3, simply because previous work in the Klobukowski group showed AM1 to be slightly superior for applications to drug-like molecules [82]. The newer semi-empirical methods, while they show promise, are not widely available at this time and were not considered. It is important to mention the consequences of using empirical data to parametrize a method such as AM1: while the results of a good method will be accurate within the set of compounds used to 'train' the method and for similar compounds, there is no guarantee that a semi-empirical method will be accurate for molecules that are different from the training set used. It may be phrased that semi-empirical methods are sufficient for interpolation (within the chemical space defined by the training set used) but extrapolation to other chemical types may not be reliable. This makes the choice of training set for new methods crucial.

It is relevant to note that AM1 and PM3 have been observed to overestimate steric repulsion and thus predict the geometry of small to medium (cyclopentane to cyclooctane size) rings to be too flat [85]. While the binary *Vinca* alkaloids studied here each contain many rings, many of these rings are aromatic or contain double bonds and are expected to be mostly flat. The artificial flatness of rings in the AM1 method is thus not expected

Table 2.1: Some calculated unsigned errors of AM1 and PM3 calculated values

Atoms	AM1	PM3
Heats of Formation (kcal mol^{-1})[a]	4.71	5.80
Bond Lengths (Å)[b]	0.027	0.022
Bond Angles (°)[b]	2.3	2.8
Dihedral Angles (°)[c]	12.5	14.9
Dipole Moment (Debye)[d]	0.35	0.38

[a]from [56], found from a set of organic compounds, some including Flourine
[b]from [83], found from a set of closed-shell organic compounds over 344 bond lengths
[b]from [83], found from a set of closed-shell organic compounds over 146 angles
[c]from [84], found from a small set (16) of data points
[d]from [77], found from a set of 125 molecules

to be a large source of error in this study.

2.4 Miscellaneous

2.4.1 Charge Fitting Methods

The fitting of partial point charges to atoms in a force field is of particular interest as the Coulombic interaction in (2.7) can contribute significantly to the potential energy. The concept of a partial atomic charge is somewhat ill-defined since it is not a fully realistic model: electrons don't reside exactly on top of atomic nuclei but rather are distributed in an electron cloud near the nuclei. It is difficult to decide which fraction of an electron cloud in a molecule 'belongs' to one atom and not to another. Therefore, the methods used to determine partial charges are varied and interesting in their own right due to the creativity often shown in their development. Several methods will be discussed here; these methods lead up to the introduction of the AM1-bcc method.

It might be expected that the assignment of charges from *ab initio* calculations would be easy, but as quantum mechanical methods model electrons as clouds, the problem

of assigning electron density to a specific atom mentioned above poses a real challenge. Mulliken [86] has proposed a simple way to assign electron density to each atom based on a Hartree-Fock calculation. Using the notation in (2.11) for the expansion of a molecular orbital over a basis set, the total number, N, of electrons in a system is:

$$\begin{aligned} N &= \sum_j \int \phi_j(j)\phi_j(j)\mathrm{d}r_j \\ &= \sum_j \sum_{\mu,\nu} \int C_{j\mu}\chi_\mu(j)C_{j\nu}\chi_\nu(j)\mathrm{d}r_j \\ &= \sum_j \sum_\mu \left(\int C_{j\mu}^2 + \sum_{\mu,\nu} C_{j\mu}C_{j\nu}S_{\mu\nu} \right) \end{aligned} \tag{2.22}$$

where j denotes a molecular orbital and $S_{\mu\nu}$ is the overlap matrix. Mulliken's scheme is to assign any electron density belonging to the first term in the last line of (2.22) to the atom that the corresponding basis set function, χ_μ, belongs to; the second sum is then seen as electron density that belongs to several atoms and this density is divided over the participating atoms. While this method has an intuitive feel to it and produces good results in certain applications it has major flaws, one of which is basis-set dependence: if the calculation is done with a different basis set, the charge assignments may be considerably different.

With the difficulties of electron density assignment illustrated above by the Mulliken charge method, it becomes apparent that another style of partial charge assignment needed to be devised. The ESP charge fitting method was pioneered by Momany [87]. The goal of the method is to reproduce the electrostatic field produced by a molecule and thus accurately model the electrostatic forces exerted and felt by the molecule. This is accomplished by calculating the SCF molecular orbital (SCF-MO) electrostatic potential on a 3-dimensional grid at approximately the molecule's Van der Waal's surface and fitting partial atomic charges by a least-squares fit of the point charge potential to the SCF-MO potential.

$$V(r) = \sum_{A,\ nuclei} \frac{Z_A}{|r - R_A|} - \sum_{\mu,\nu} P_{\mu\nu} \int \frac{\chi_\mu \chi_\nu}{|r - r'|}\mathrm{d}r' \Longrightarrow \sum_{j,\ atoms} \frac{q_j}{\left|r - r'_j\right|} \tag{2.23}$$

The point charges will then produce a potential around the molecule that is very close to the SCF-MO one. Momany's original paper used experimental dipole moments in the

determination of the point charges and the method was shown to effectively reproduce the electrostatic potential around a molecule. Further testing and refinement of the method was done by Cox [88] who also showed the method to be effective. Cox's method did not use experimental dipole moments but rather only the calculated SCF-MO potentials. As it was designed to, the method produces atomic charges that model the electrostatics of a molecule effectively, and thus makes the method well suited for force field molecular modeling. However, there are also potential problems with the ESP method. The charges were found to be basis-set dependent for small basis sets up to, but stable above, the 6-31G** level [88] and thus require a reasonably high-level basis set to be accurate. The method is also dependent on the conformation of the molecule as intramolecular effects can polarize a molecule in different ways depending on its conformation. The largest flaw in ESP-fitting is that the method can produce unreasonably large charges for atoms that are 'buried' [58], such as a methyl carbon atom. This occurs since the least-squares fit matches the potential outside of the molecule and this potential will be less affected by such buried atoms than by surface atoms due to proximity. Thus, the charge on buried atoms can be 'poorly determined' by the method and can become large.

The ESP method was modified by Bayly [58] to address the problem of unreasonable buried charges. A restraint was used in the fitting procedure to restrain the atomic charges to zero. This approach was a success since it produced more reasonable charges of buried atoms while having little effect on surface atoms. This method, Restrained Electrostatic Potential fitting (RESP), has become a standard in production of atomic charges for force fields [89] and is the method used for the popular AMBER force fields [90, 91]. While RESP is a large improvement over ESP, it still requires time consuming high-level *ab initio* calculations, so while it is appropriate for finding parameters that will find repeated use such an amino acid force field, it is not suitable for producing parameters for specific molecules that will only be used once. Like ESP, RESP is conformationally dependent: this may be improved upon by the use of multi-configurational or ensemble RESP fitting, which factor in several viable conformations of a molecule. However, these methods then drastically increase the number of SCF calculations required. A problem with RESP fitting at the 6-31G** level is that it overpolarizes molecules [58]. This overpolarization is serendipitous, though, since it matches the overpolarization found in the TIP3P water model, and thus makes RESP charges appropriate to reproduce solvent interactions with the TIP3P model.

The AM1-bcc [92, 93] method involves applying parametrized bond charge corrections

(bcc) of low-level Mulliken charges from an AM1 calculation. The charge on each atom is given by:

$$q_j = q_j^{AM1} + q_j^{CORR} \tag{2.24}$$

where q_j^{AM1} is the Mulliken atomic charge for atom j and q_j^{CORR} is a correction for each atom based on its bonding environment. The essence of the method is that the AM1-Mulliken charges roughly capture the chemical nature of the molecule in question, and then they are corrected to achieve a higher level of accuracy. The q_j^{CORR} term is a linear sum of correction parameters; each parameter depends on the type of bond (single, double, triple, aromatic) and the types of atoms involved in the bond:

$$q_j^{CORR} = \sum_{\alpha} T_{j\alpha} p_{\alpha} \tag{2.25}$$

were T is a matrix defining connectivity of the atom j of each bond-type α and p_α is the correction term for bond type α. The corrections are applied in a conservative manner: charge is transferred from one atom and placed on a bonded atom so as to conserve the net charge of the molecule. The bond charge correction parameters were parameterized using a set of over 2700 representative organic molecules to reproduce RESP 6-31G* calculations and have been found to do so quite well, with a correlation coefficient of over 0.95. As the method is much faster than RESP, it is a valid alternative for fitting specific molecules outside of standard force fields.

2.4.2 The QM/MM Hybrid Method

The two theoretical formalisms discussed so far in this chapter, Quantum Mechanical (QM) and the classical Molecular Mechanics (MM) methods, appear to be opposite and mutually disjoint: quantum mechanics handles electronic properties and excited states but is very slow and limited in the size of a system it can handle; standard molecular mechanics is only good for systems near equilibrium but is very fast and can handle large systems. While the semi-empirical quantum mechanical methods also described are much faster than full *ab initio* methods, they are still not as fast as force field molecular mechanical methods and are not able to model large systems such as proteins. A middle-ground approach to theoretical modeling is a hybrid of the QM and MM formalisms: the QM/MM method. The goal behind such a method is to enable the study of a large system described mostly by the efficient molecular mechanics method and to have a relatively small portion described by the more robust quantum mechanics method.

A standard application of the QM/MM method is the study of an enzyme, where the majority of the system (protein) is well-described by molecular mechanics, but a small active site, which may undergo bonding changes or at the very least require explicit electron density to correctly model the enzymatic action, is described by quantum mechanics. Another desirable application, as used in the current study, is the modeling of a solute by quantum mechanics in an effort to capture a more accurate description of the molecule and its electronic properties but with explicit molecular mechanical water molecules, using the TIP3P model for example, to include solvation.

The Hamiltonian of a QM/MM system is:

$$H = H_{QM} + H_{MM} + H_{QM/MM} \tag{2.26}$$

where H_{QM} describes only the quantum mechanical region, H_{MM} only the molecular mechanical region, and $H_{QM/MM}$ describes the interaction between the two regions. There are two possibilities for the interaction Hamiltonian: 1) there are no chemical bonds between the two regions; and 2) there are chemical bonds between the two regions. The first case is much simpler than the second. There are no chemically-bonded interactions in the use of QM/MM methods in this thesis and thus the following discussion is limited to this case. The interaction Hamiltonian is then limited to electrostatic and dispersion forces:

$$H_{QM/MM} = \sum_{i}^{\substack{QM \\ electrons}} \sum_{m}^{\substack{MM \\ atoms}} \frac{q_m}{r_{im}} + \sum_{k}^{\substack{QM \\ nuclei}} \sum_{m}^{\substack{MM \\ atoms}} \frac{Z_k q_m}{r_{km}} + \left(\frac{a_{km}}{r_{km}^{12}} - \frac{b_{km}}{r_{km}^{6}} \right). \tag{2.27}$$

The above equation is divided into two sums, the first involves electrons and will then act upon an electronic wavefunction, and the second will not operate on an electronic wavefunction. The consequence of this separation is that, using the Dirac bra-ket notation[1], the expectation value of the energy is:

$$\langle \Psi | H | \Psi \rangle = \left\langle \Psi \left| H_{QM}^{*} \right| \Psi \right\rangle + H_{MM}^{*} \tag{2.28}$$

where all of the electron containing terms are in H_{QM}^{*} which can then be solved using the SCF procedure described for the Hartree-Fock method and all other terms, which are purely classical in nature, are in H_{MM}^{*}. H_{QM}^{*} is a hamiltonian that is very similar to

[1]Dirac's notation states that a bra, $\langle \Psi |$, and a ket, $| \Psi \rangle$, when used together, act as an integral: $\langle \Psi | H | \Psi \rangle = \int \Psi^{*} H \Psi d\tau$

the pure quantum mechanical hamiltonian, H_{QM}, and H^*_{MM} is of very similar form to the pure force field term H_{MM}. The difference in the quantum mechanical hamiltonian is that any sum over quantum mechanical nuclei now includes the molecular mechanical atoms: the quantum mechanical nuclei have full nuclear charges, Z_k and the molecular mechanical atoms have their partial atomic charges, q_k as assigned by the force field used. The effect of the inclusion of molecular mechanical atoms in this way is seen in the Fock matrix, which becomes

$$\mathbf{F}_{\mu\nu} = \int \chi_\mu \left(-\frac{1}{2}\nabla^2\right) \chi_\nu \mathrm{d}r - \sum_k^{\substack{QM \\ nuclei}} Z_k \int \chi_\mu \left(\frac{1}{r_k}\right) \chi_\nu \mathrm{d}r$$
$$- \sum_k^{\substack{MM \\ atoms}} q_k \int \chi_\mu \left(\frac{1}{r_k}\right) \chi_\nu \mathrm{d}r + \sum_{\lambda,\kappa} P_{\lambda\kappa} \left[\int \chi_\mu(1)\chi_\nu(1) \frac{1}{r_{12}} \chi_\lambda(2)\chi_\kappa(2) \mathrm{d}r_1 \mathrm{d}r_2 \right.$$
$$\left. - \frac{1}{2} \int \chi_\mu(1)\chi_\lambda(1) \frac{1}{r_{12}} \chi_\nu(2)\chi_\kappa(2) \mathrm{d}r_1 \mathrm{d}r_2 \right] \qquad (2.29)$$

where the second term in the standard Fock matrix (2.13) is simply repeated with another sum over molecular mechanical partial atomic charges. The difference in the molecular mechanical hamiltonian is that now the quantum mechanical nuclei are included in the pair wise interaction sums as described in (2.7) and (2.8). The dispersion forces across the QM/MM boundary are then described by a Leonard-Jones interaction, just like between two molecular mechanical atoms. This is important because hydrogen bonding between a quantum mechanical atom and an molecular mechanical atom (in this case a solute atom and a solvent TIP3P atom, respectively) will be handled in an accepted manner. The modified Hartree-Fock equations are then solved with the usual SCF method.

2.4.3 Energy Minimization and Molecular Dynamics

In the above text, three methods of determining the potential energy of a chemical system have been discussed. The question remains, however, what to do with this knowledge of potential energy? Before the potential energy can be found, though, what does the system look like? The entire set of atomic coordinates, the geometry of the system, must be known before any of the methods above can be used. Experimentally determined geometries could be used as input, but that would limit theoretical chemistry to systems where high-resolution experimental geometries are available. A much more powerful use of theoretical chemistry would be in the prediction of geometries and other properties.

Thus, if a minimum energy geometry of a chemical system can be found, this geometry will be stable and be a reflection of reality. If a minimum energy geometry is found, the spatial gradient of the potential energy will be zero, and thus from (2.1) the force will be zero. Since an object will move in the direction of decrease of potential energy, one obvious method of obtaining a minimum geometry is to allow the system to move with the forces on the system until the forces are all zero. Of course, there are more sophisticated methods available, but these will not be discussed. The process of finding an energy minimum is called energy minimization or geometry optimization.

Energy minimization is a valuable tool of computational chemistry and is used extensively in the current work. However, the result of such a calculation is a static structure with no atomic motion. This means that the structure is representative of a system at 0 K since atomic motion is proportional to thermodynamic temperature. Couldn't a different description of the system involve temperature and thus atomic motion? The answer is yes, and this method is called Molecular Dynamics (MD). In molecular dynamics, each atom is given velocity to produce a temperature T, forces are calculated using one of the methods described above, and then the system is moved through finite time-steps allowing for a dynamic evolution of the system. There are many intricacies involved in MD simulations and they will not be described in detail here. Perhaps the most important is a computational thermostat: thermostats can be used to control the temperature (in other words, the kinetic energy) of the system. Without a thermostat, the temperature may decrease or increase to an undesired value. Over the time evolution of the system, a number of properties can be ascertained using a variety of techniques: geometries and average energies, enthalpies, and free energies are some of the most important.

2.5 The Blocking Method

MD simulations produce a series of data points, taken as 'snapshots' at discrete time points. These data points are not independent, but are *correlated*. As a result, statistical measures designed for use with correlated data must be used. One such method is the Blocking Method. While the original source of the blocking method is not known, Flyvberg [94] has described its use – for a thorough description of the method the reader is encouraged to read this article. The method produces an estimate of the error of the data set, and additionally calculates a standard deviation of the error. The error standard deviation calculation is important as it allows determination of whether or not

the method has converged to a valid result: if the standard deviation is too large, the simulation is not long enough or the measured parameter is too unstable and the error should not be trusted. Most often, an increased simulation length will reduce the error standard deviation.

Chapter 3

Free *Vinca* Alkaloids

The overall action of the *Vinca* alkaloid drugs on the human body involves many complicated processes, including distribution within the body, metabolism, and ligand-protein interaction, to name a few. This can lead to a highly complex system and will likely not be fully understood for a long time to come. While the ultimate interest in pharmaceuticals must be their action in the human body, there is merit in studying the drugs in an isolated state as their individual properties are an integral part of their action. The information obtained about the drugs themselves can then be used to help develop models of action, and help explain their pharmacological profiles. In particular, because minor changes in the *Vinca* alkaloids studied in this thesis give rise to significant differences in the pharmacological profiles of the drugs, comparative studies between these five *Vinca* alkaloids may prove fruitful. While a whole host of data are available from computational studies, these studies are particularly interesting in that they will provide information about what these drugs look like: computational methods can provide accurate atomic resolution geometries of molecules, and while seemingly trivial this can be powerful information. In the case of the *Vinca* alkaloids, high-resolution geometric information is very important to obtain, because such experimental data of the drugs in environments similar to that of the human body is nearly non-existent. The computational work presented in this section will attempt to provide atomic-level information of the *Vinca* alkaloids in aqueous solution at physiological temperature.

3.1 Torsional Potential Energy Surfaces About the Main Dihedral Angle

Inspection of the structure of the *Vinca* alkaloids reveals that the main (C17-C18-C15-C16) dihedral angle is of interest: it connects two relatively large groups and torsion about this bond will result in different molecular shapes. This is relevant for a number of reasons, but perhaps the most important is, as both top and bottom portions of the *Vinca* alkaloids are apparently involved in binding to tubulin, that the binding interfacial regions may be different at different main dihedral angles. The projection of the potential energy surface at the semi-empirical AM1 level of each *Vinca* alkaloid onto the main dihedral angle has been calculated in an effort to understand the rotational properties about this angle.

3.1.1 Procedure

Starting coordinates for VLB were taken from Bau and Jin's crystal structure [1]. This choice will ensure correct stereochemistry of VLB. These coordinates were inputted into the SPARTAN [95] molecular building environment on an Apple PowerMac G5 Quad 1.9 GHz computer. Structures of VCR, VDE, VNO, and VFL were created by mutating the VLB structure at the required functional groups and in addition by deleting the C8' atom in the latter two cases. Each structure was then optimized at the MMFF [96–100] level in the SPARTAN building environment. For each drug, 72 models were created: one for each of the angles between 0° and 355° at 5° increments. The main dihedral angle was restrained at each angle, and a Monte Carlo conformer search was performed at the MMFF level for each dihedral angle. A Monte Carlo search was not performed on VFL as the SPARTAN program was found to erroneously reverse the stereochemical configuration at the C18' position during the search – instead, the structures of VFL were manually prepared based on the results from the other four drugs. The results of such a search will give a single structure, under the restraints specified, that will be a 'good guess' at a global minimum structure. The coordinates resulting from the Monte Carlo search were then inputted into PC-GAMESS [74], and keeping the main dihedral restrained at the same 5° increments, *in vacuo* geometry optimizations were performed at the semi-empirical AM1 level. The PC-GAMESS calculations were performed on the OH cluster (comprised of AMD Athlon 2.0GHz computers). Restraints were imposed using the `$zmat` keywords `IFZMAT`, `FVALUE`, and by setting `AUTOFV=.false.`, which is a feature unique to PC-GAMESS in the GAMESS family of programs. Although molecular geometries were inputted as cartesian coordinates, the calculations used automatically

generated delocalized internal coordinates. The Hessian matrix was calculated at the initial optimization step, and then subsequently every 10 steps afterwards, with the BFGS Hessian update method [101] performed at each step in between. SCF density convergence of less than 10^{-5} was required, and in the Rational Function Optimization (RFO) [102] routine a geometry convergence criteria of 10^{-4} hartree bohr^{-1} (for the largest component of the gradient) was used. The energies of the optimized structures were converted to units of kcal mol^{-1} plotted against the dihedral angle and this is the projection of the potential energy surface on the torsion of the main dihedral angle. The graphs of each of the five drugs displayed 'spikes' and 'plateaus' of unexpectedly high energy. It is likely that these high energy regions were the result of the systems becoming caught in local minima. To remedy this, the structures from neighboring low energy angles were shifted in their main dihedral angles by five or ten degrees to the angles of the high energy structures. These shifted structures were then optimized again. This procedure allowed for the lowering of most unusually high energy regions, leaving a smoother surface. Some minor spikes are still present and could not be removed.

3.1.2 Results

The five potential energy surface plots are presented in Figs. 3.1 to 3.5.

Of note are the minor spikes in the 0-100° region of the VLB surface and the spikes near 180° in each of VLB, VCR, and VDE. The plots for VLB, VCR, and VDE are all very similar, each with a global minima at about 205° and with secondary minima at about 35°, as well. The relative depth of the secondary minima are about 2.5 kcal mol^{-1} in both VLB and VDE, but the depth of this minimum in VCR is about 0.5 kcal mol^{-1}. The VNO and VFL plots are similar to the other three, but with a global minimum at 195° and a secondary minimum at about 20° which is 3 kcal mol^{-1} higher in energy than the global minimum.

As the main dihedral angle is changed about a full rotation, structural changes occur. While minor conformational changes, such as torsion of a small functional group, are visible throughout the molecules, major rearrangements occur in the backbone structure of the top portion. At the minimum energy conformations near 200° and 40° the top portion closely resembles the geometry found by Bau and Jin [1]. In the 300-360° domain, the top portion adopts a conformation that is considerably bent at the middle. The structure at the energy plateau near 100° is like the bent conformation near 300° but can become very distorted and even the chair conformation of the piperidine ring is lost. Each of the drugs undergo these same conformational changes, including VNO and

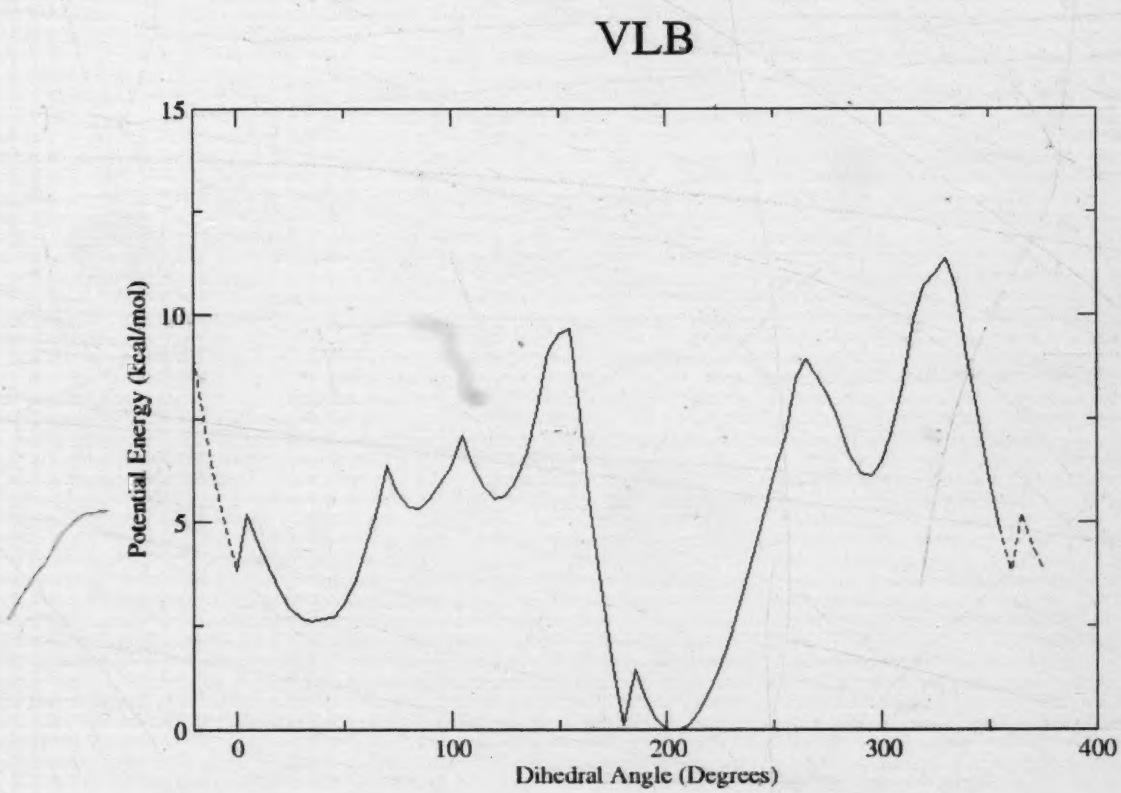

Figure 3.1: The projection of the potential energy surface of VLB onto the C17-C18-C15-C16 dihedral angle. The graph is periodically wrapped in the -20-0° and 360-380° domains (in dotted lines) for increased clarity.

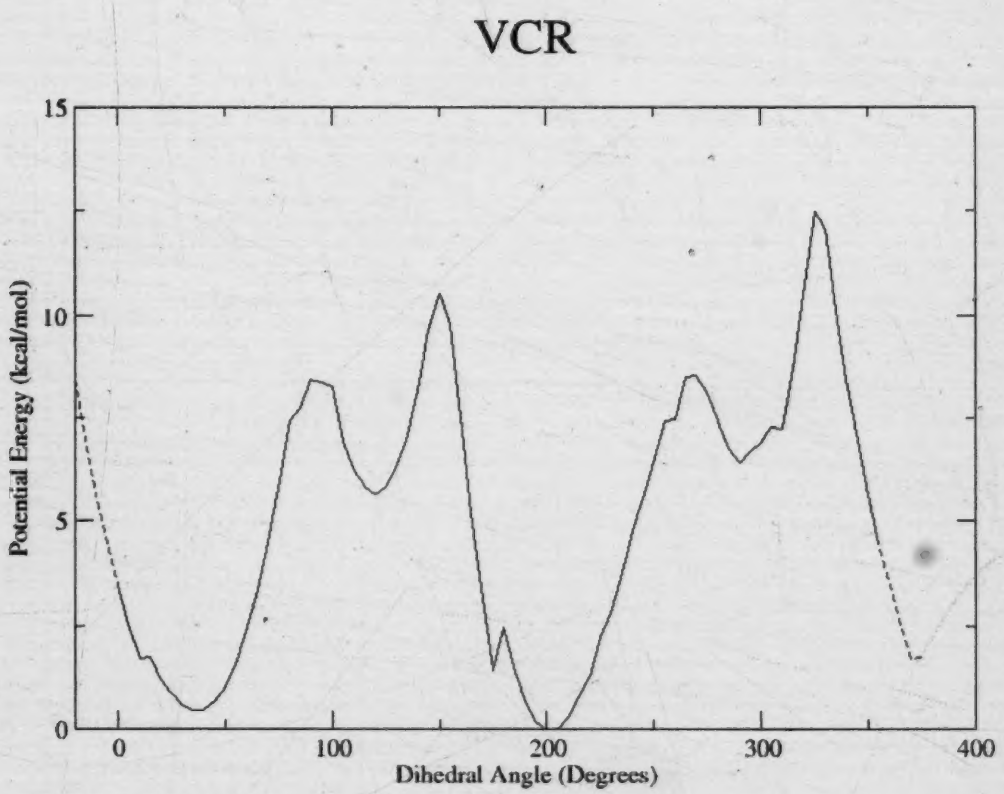

Figure 3.2: The projection of the potential energy surface of VCR onto the C17-C18-C15-C16 dihedral angle. The graph is periodically wrapped in the -20-0° and 360-380° domains (in dotted lines) for increased clarity.

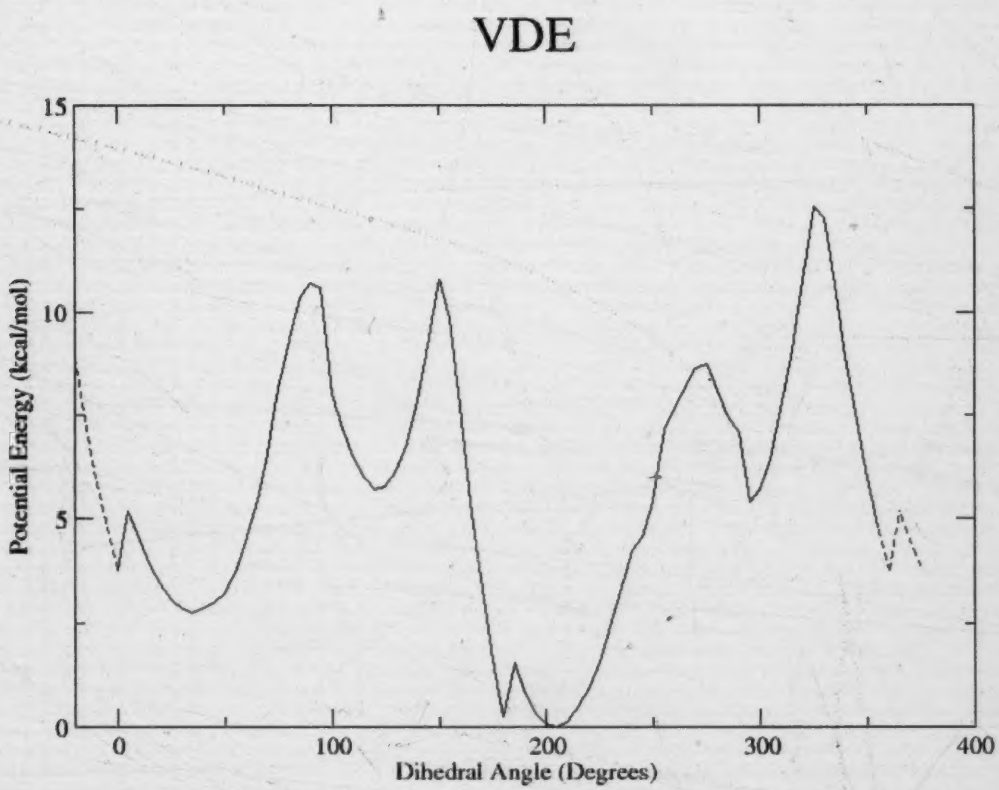

Figure 3.3: The projection of the potential energy surface of VDE onto the C17-C18-C15-C16 dihedral angle. The graph is periodically wrapped in the -20-0° and 360-380° domains (in dotted lines) for increased clarity.

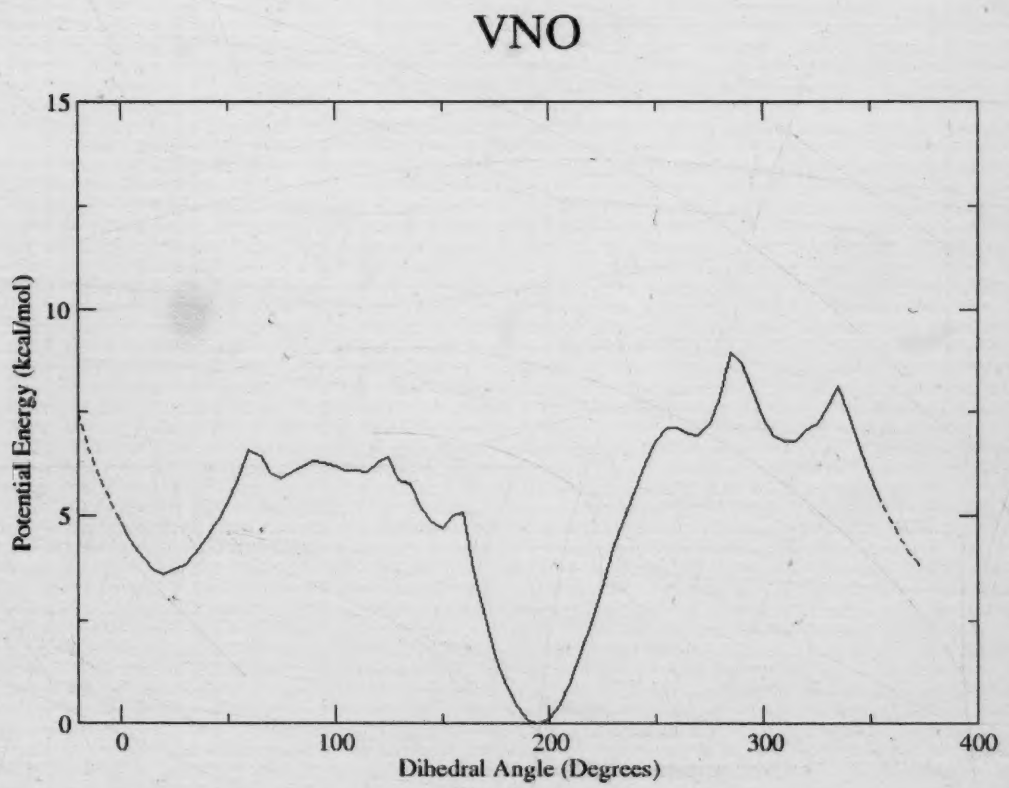

Figure 3.4: The projection of the potential energy surface of VNO onto the C17-C18-C15-C16 dihedral angle. The graph is periodically wrapped in the -20-0° and 360-380° domains (in dotted lines) for increased clarity.

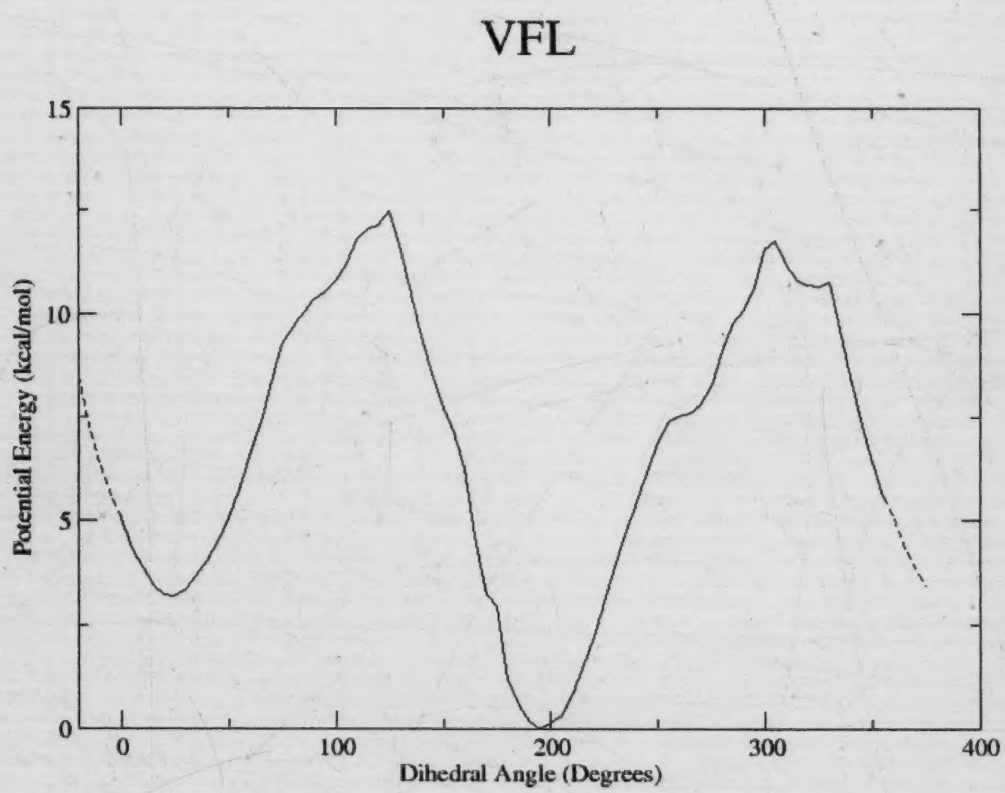

Figure 3.5: The projection of the potential energy surface of VFL onto the C17-C18-C15-C16 dihedral angle. The graph is periodically wrapped in the -20-0° and 360-380° domains (in dotted lines) for increased clarity.

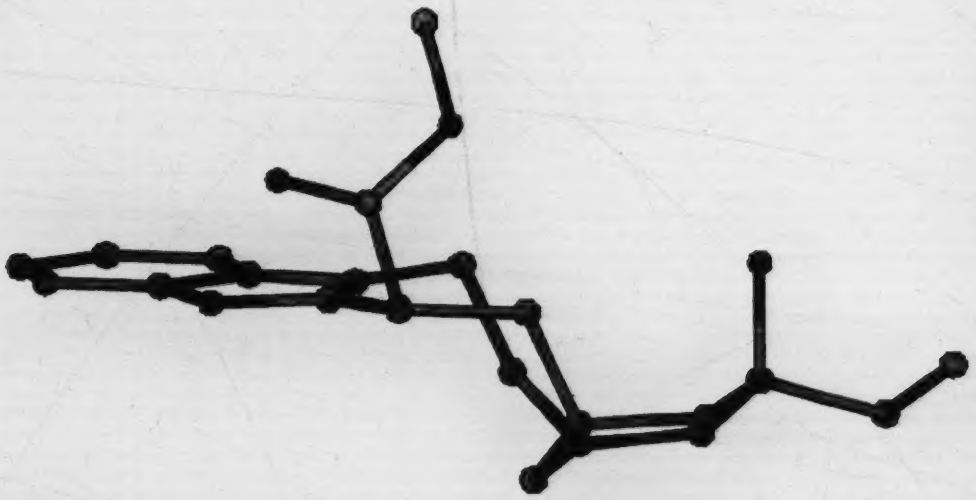

Figure 3.6: Top portion of VLB from the AM1 optimized 200° main dihedral angle structure. Hydrogen atoms are removed for clarity. This structure is similar to the one found using X-ray crystallography by Bau and Jin [1] in that the piperidine ring is in a chair conformation and the C18', C1', and C8' atoms are nearly coplanar with the indole moeity. VCR, VDE, and VFL adopt similar conformations near minima of their torsional potential energy surfaces. Figure created using VMD [31].

VFL, despite the smaller and less flexible 8-membered ring and the double bond in the piperidine ring of VNO. Of course, the double bond in the piperidine ring of VNO forces that 6-membered ring into a near-planar shape. Using VLB as an illustrative example, Figs. 3.6 to 3.8 show the varying conformations of the top portions of the drugs. Several VNO and VFL conformations are shown in Figs. 3.9 to 3.11.

The dipole moments calculated in the *in vacuo* AM1 energy minimizations, at several conformations each, of the *Vinca* alkaloids are presented in Table 3.1.

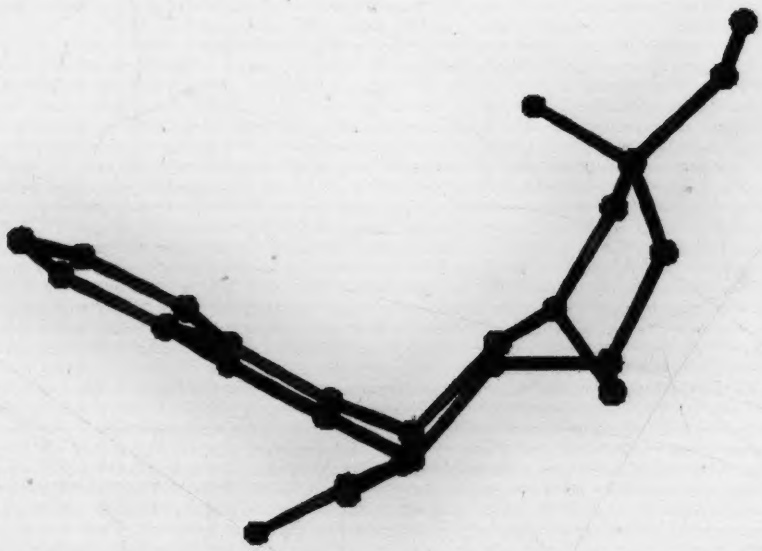

Figure 3.7: Top portion of VLB from the AM1 optimized 350° main dihedral angle structure. Hydrogen atoms are removed for clarity. This structure differs from the one found using X-ray crystallography by Bau and Jin [1] in that the coplanarity of the indole moeity and C18' is lost, but the chair conformation of the piperidine ring is preserved. VCR and VDE adopt similar conformations near peaks of their potential energy surfaces. Figure created using VMD [31].

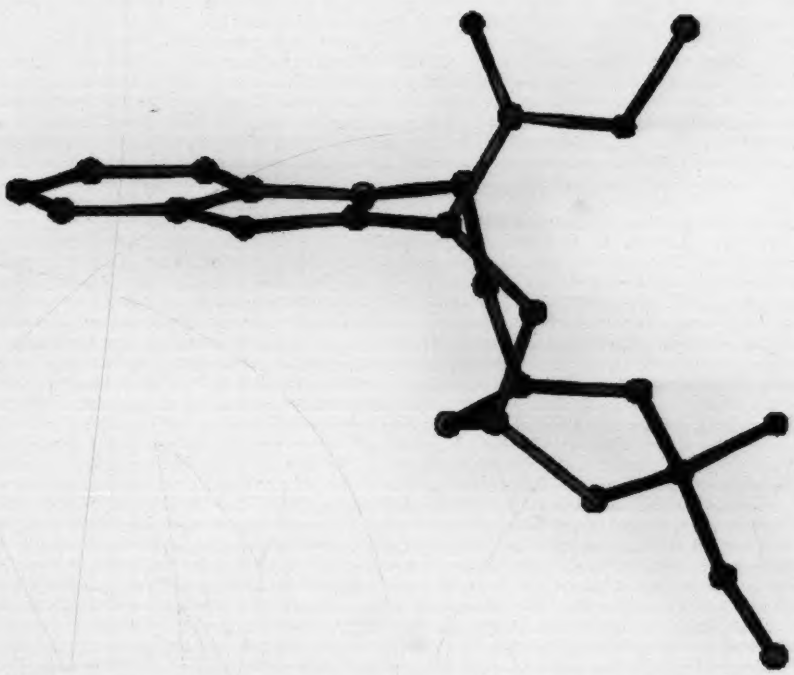

Figure 3.8: Top portion of VLB from the AM1 optimized 110° main dihedral angle structure, as an example of a very distorted, high energy structure of VLB. Hydrogen atoms are removed for clarity. This structure is significantly different from the one found using X-ray crystallography by Bau and Jin [1] in that the coplanarity of the indole moeity and C18' is lost and the piperidine ring is distorted out of the chair conformation. Most other high energy structures of the *Vinca* alkaloids adopt a bent conformation. Figure created using VMD [31].

Table 3.1: *in vacuo* AM1 Dipole Moments at the Minima of the *Vinca* Alkaloids. 'Angle' refers to the C17'-C18'-C15-C16 dihedral angle.

Drug	Angle(°)	Dipole Moment (D)
VLB	205.0	2.137515
	180.0	2.672800
	195.0	1.881025
	215.0	2.215069
	35.0	2.878782
VCR	200.0	3.601623
	175.0	5.813018
	190.0	3.642041
	215.0	3.483965
	35.0	5.428539
VDE	205.0	2.877218
	180.0	4.311632
	195.0	2.753352
	215.0	2.873112
	35.0	2.981671
VNO	195.0	2.360743
	185.0	2.233963
	205.0	2.438230
	20.0	3.323205
VFL	195.0	2.687364
	185.0	2.421758
	205.0	2.992591
	20.0	5.846031

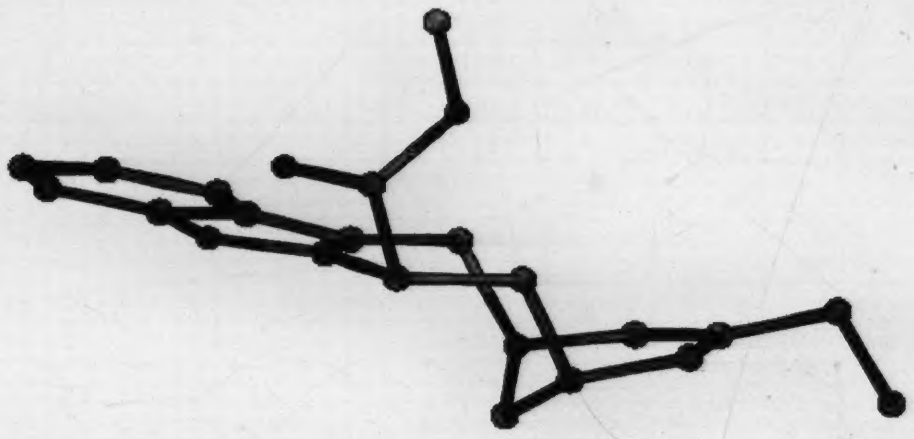

Figure 3.9: Top portion of VNO from the AM1 optimized 195° main dihedral angle structure, as an example of a low energy, stable conformation of VNO. Hydrogen atoms are removed for clarity. Most structural features found in the VLB crystal structure of Bau and Jin are preserved here. Figure created using VMD [31].

3.2 QM/MD simulations

The torsional potential energy surfaces presented above, as the result of *in vacuo* geometry optimizations, are valuable because they readily give a quantifiable understanding of how the *Vinca* alkaloids will behave. However, *in vacuo* geometry optimizations are not necessarily the most realistic representation of a pharmaceutical molecule because: 1) pharmaceutical compounds are active in the human body, and their environment (in this case aqueous) must be accounted for; and 2) a geometry optimization is a 0 K representation and pharmaceutical compounds are active at body temperature, which is approximately 310 K. Thus, the potential energy surfaces presented above are useful in understanding the *Vinca* alkaloids, but a more robust method should be looked for. To address these two inadequacies in the geometry optimization method, a molecular dynamics simulation utilizing a water model to include solvation is a natural choice. The exact same semi-empirical AM1 theory used in the geometry optimizations can be applied to the solute drug molecule by using the hybrid QM/MM theory along with the TIP3P water model. Molecular dynamics simulations based on QM/MM theory are referred to as QM/MD for short.

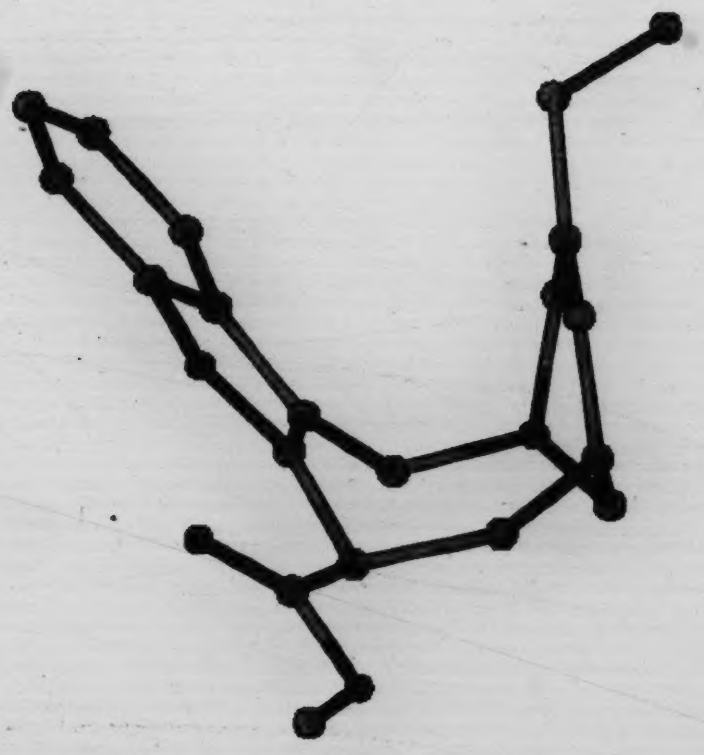

Figure 3.10: Top portion of VNO from the AM1 optimized 300° main dihedral angle structure, as an example of a bent, high energy structure of VNO. The top portion of VNO takes on a more extreme bend at these high energy structures when compared to VLB, VCR, and VDE. Hydrogen atoms are removed for clarity. Figure created using VMD [31].

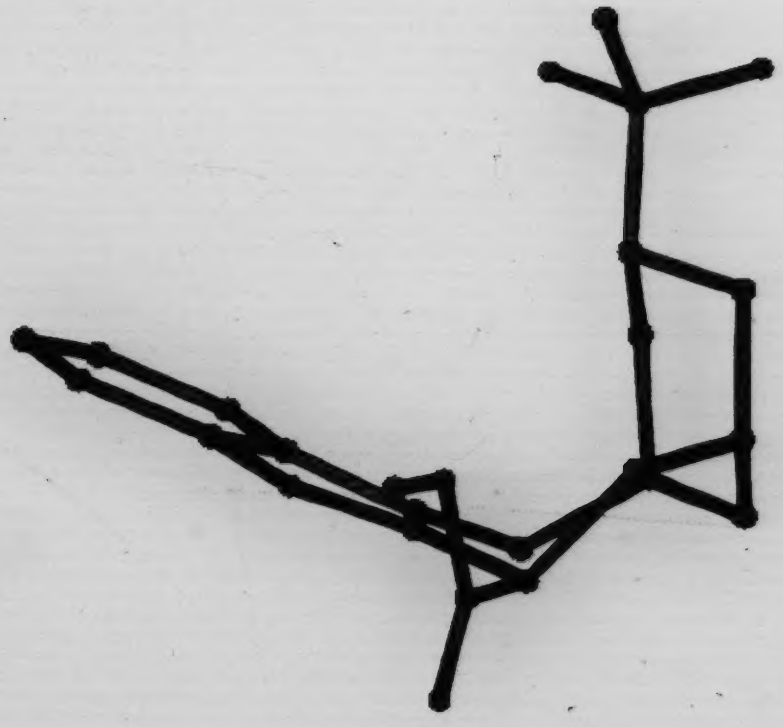

Figure 3.11: Top portion of VFL from the AM1 optimized 325° main dihedral angle structure, as an example of a bent, high energy structure of VFL. The top portion of VFL takes on a more extreme bend at these high energy structures when compared to VLB, VCR, and VDE. Hydrogen atoms are removed for clarity. Figure created using VMD [31].

3.2.1 Procedure

Again using the SPARTAN program, the same starting structures of each drug used in the geometry optimizations were used. An unrestrained Monte Carlo search was performed at the MMFF level on each drug, and the 20 structures with the lowest energy were analyzed. The resulting structures were binned by main dihedral angle in groups of 10° width and the lowest energy structure from each bin was chosen to be used in an MD simulation. In addition, several angles were chosen from the geometry optimized structures in a effort to better represent certain areas of the torsional potential energy surface. For VFL, no Monte Carlo search was performed due to the tendancy of a SPARTAN Monte Carlo search to reverse the stereochemistry at the C18' site. All of the structures chosen for VFL were taken from its torsional potential energy surface structures. As there are multiple structures simulated for each drug, the notation of VLB_N, where N is an index to describe the starting conformation, will be adopted when discussing these QM/MD simulations. The main dihedral angles of the starting structures are presented in Table 3.2. The `solvatebox` command from the TELEAP program of the AMBER [75] suite was used to solvate each drug molecule by periodically placing pre-equilibrated boxes of TIP3P water around the solute drug. Water molecules that were placed within the Van der Waals radius of any solute atom were automatically removed. Exactly 2000 water molecules were added to each simulation box.

Born-Oppenheimer QM/MD simulations were performed using the SANDER program from the AMBER 9 suite [75]. The leap-frog integrator was used to propagate the system in molecular dynamics. QM forces were calculated using an SCF procedure. An SCF cycle was considered converged when an energy convergence criterion of 10^{-8} kcal mol^{-1} and a largest element-wise change in density matrix criterion of $5 * 10^{-10}$ were both met. An electrostatic cutoff of 10 Å was used for the MM region and a cutoff of 8 Å was used for the QM region. It is important to note that these two cutoffs have different meanings: within the MM region, two atoms separated by more than the 10 Å cutoff will not be included in the non-bonded interaction energy; there is no non-bonded cutoff between QM atoms, but rather the 8 Å cutoff is from the edge of QM/MM boundary out into the MM region. Thus, any MM atom that is more than 8 Å away from the boundary is not included in the QM electrostatics (see equation (2.29)). A Particle Mesh Ewald method (PME) modified for QM/MM systems [103] was applied to long range electrostatics. The Langevin thermostat was used in MD simulations to control the temperature. The SHAKE [104] algorithm was used to constrain bond lengths: this was applied only to bonds involving hydrogen and all such bonds (solute and water) were

Table 3.2: The starting conformations of the *Vinca* alkaloids in QM/MD simulations. 'Angle' refers to the C17'-C18'-C15-C16 dihedral angle.

Drug and Index	Angle(°)	Source
VLB_1	205.3	Monte Carlo Search
VLB_2	154.9	Monte Carlo Search
VLB_3	311.3	Monte Carlo Search
VLB_4	37.6	Monte Carlo Search
VLB_5	175.0	Torsional PES
VLB_6	120.0	Torsional PES
VCR_1	153.1	Monte Carlo Search
VCR_2	202.0	Monte Carlo Search
VCR_3	311.9	Monte Carlo Search
VCR_4	165.4	Monte Carlo Search
VCR_5	36.3	Monte Carlo Search
VDE_1	205.0	Monte Carlo Search
VDE_2	155.2	Monte Carlo Search
VDE_3	311.1	Monte Carlo Search
VDE_4	168.2	Monte Carlo Search
VDE_5	37.5	Monte Carlo Search
VDE_6	316.4	Monte Carlo Search
VDE_7	120.0	Torsional PES
VNO_1	195.4	Monte Carlo Search
VNO_2	11.1	Monte Carlo Search
VNO_3	315.0	Torsional PES
VFL_1	195.0	Torsional PES
VFL_2	20.0	Torsional PES
VFL_3	285.0	Torsional PES
VFL_4	100.0	Torsional PES

restrained to equilibrium bond lengths at all times during MD simulations.

Before MD simulations were started brief and incomplete geometry optimizations were performed in order to remove any large forces due to potentially unusual geometries in the starting structure that might cause extreme behaviour or crashes on starting an MD simulations. Using the SANDER program's energy minimization function, on an Apple iMac G5 1.9 GHz computer, incomplete optimizations were performed. In an effort to speed up the calculation, the solute was represented with the Generalized Amber Force Field (GAFF) [105] and AM1-bcc charges[1] Periodic box boundary conditions were used with the box volume held constant. In a first step, the solute atoms were all restrained to their original positions using parabolic energy penalties[2] with a force constant of 100 kcal/mol/Å^2. The steepest descent algorithm was used for 250 iterations, and then the conjugate gradient method was used for 750 iterations. A second step was subsequently performed, with no positional restraints on the solute, for 500 iterations of steepest descent and then 2500 steps of conjugate gradient minimization.

The solvent was then 'warmed' by performing a short molecular dynamics simulation with parabolic restraints on the solute. The force constants used were 20 kcal/mol/Å^2, and the solute was restrained to the coordinates resulting from the previous two stages of energy minimization. The Langevin thermostat was used with a collision frequency of 5 ns^{-1} and a target temperature of 310 K. An integrator time step of 0.5 fs was used and 10000 steps of molecular dynamics simulation was performed, under periodic box conditions with a constant volume. Next, the exact conditions (solute restraints, periodic box conditions, thermostat) were used for solvent equilibration with an integrator time step of 2fs over 25000 simulation steps to a total of 50 ps. The same GAFF/AM1-bcc description of the drugs was applied here in an effort to speed up the process[3]. The warming and solvent equilibration stages were all performed on an Apple iMac G5 1.9

[1]The force-field representation of solute was used mostly with the solute restrained and therefore no geometric differences would result compared to using the slower AM1 method. The GAFF representation was only used unrestrained for a small number of energy minimization steps, therefore the resulting geometry would not be appreciably different than if an AM1 representation had been used. Furthermore, the GAFF force-field was tested and proved to not accurately model the *Vinca* alkaloids, so unrestrained simulation using this force-field would not be advisable.

[2]A parabolic energy penalty is an energy penalty of the form $U_{restraint} = \frac{1}{2}k(x - x_o)^2$ where k is the force constant, x is a positional coordinate, and x_o is the value of the coordinate to restrain to. This results in a force of $F_{restraint} = -k(x - x_o)$.

[3]Caution must be exercised when using a force-field method to equilibrate a system for QM/MD simulations as the two methods may describe the system differently and conformational changes may occur during equilibration that would not occur in the molecular dynamics simulation. GAFF/AM1-bcc was used for dynamics simulations only when the solute was restrained.

GHz computer using the SANDER program.

Next, the entire system was equilibrated, without restraints, and at the AM1/TIP3P solute/solvent level of theory. Periodic box conditions were again used, only this time with the use of a barostat to yield a constant pressure of 1 bar. A pressure relaxation time of 2 ps^{-1} and isotropic box scaling were the parameters used for pressure regulation. The collision frequency for the Langevin thermostat was set to 1 ps^{-1} and a temperature of 310 K. A time-step of 2 fs was used over 50000 steps to give a total equilibration time of 100 ps. Each system was judged to be equilibrated once the potential energy, kinetic energy, pressure, and density of the system reached stable values with no increasing or decreasing drift. This point was typically obtained by the 50 ps mark of the unrestrained equilibration simulation, but an extra 50 ps was performed to err on the side of caution. The exact same conditions were used for the production simulations, which were performed for a total of 5 million time steps to give 1 ns of simulation. The full equilibration and production runs were performed on the Glacier cluster (3.06 GHz Intel Xenon processors) of the Westgrid high performance computing consortium. Data points were taken every 0.1ps and analysis of the simulations was performed with the PROCESS_MDOUT.PERL script and PTRAJ utilities from the AMBER suite of programs. Radii of gyration were calculated by the usual formula. In addition to the main dihedral angle, the linear distance between the C4' atom on the top portion of the drugs and C10 atom of the bottom of the drugs were calculated. This distance was chosen as an alternative, structure independent, geometric parameter to the main dihedral angle. Both of these atoms are predicted by the 1Z2B structure to be in contact with tubulin in a bound state, thus their relative position may be important. Statistical errors were calculated using the blocking method [94].

3.2.2 Results

Plots of the C17'-C18'-C15-C16 main dihedral angles as they vary over time are presented in Figs. (3.12) to (3.36). The first 50 ps of each plot are from the solvent equilibration stage where the solute drug is restrained – this region of each plot will show the starting value of the main dihedral angle for that simulation. The next 100ps are the QM/MD equilibration and the remaining 1 ns is the production QM/MD simulation. One simulation for each drug is started very near one of the two minimum energy main dihedral angles found in the 0 K AM1 torsional potential energy surfaces. These simulations remain stable at the starting value of the main dihedral, indicating that the 0 K minimum energy angles are also the equilibrium angles in finite temperature simulation. The other

simulations exhibit two types of behaviour: 1) they collapse directly to an equilibrium main dihedral angle; or 2) they eventually collapse to one of the two equilibrium angles, after staying in one or more metastable angles for up to 0.5 ns or so. The most important thing to note about these plots is that, despite initial variance in structure, each molecule will eventually adopt one of two values of the main dihedral. Thus, the results from the simulations started from each of the equilibrium values are the most important. The averaged results of these simulations are presented in Table (3.3) along with the C4'-C10 distances. Also in Table (3.3) are averaged radii of gyration, AM1 energies of the solute drug, and overall QM/MM energies of the entire drug and solvent system.

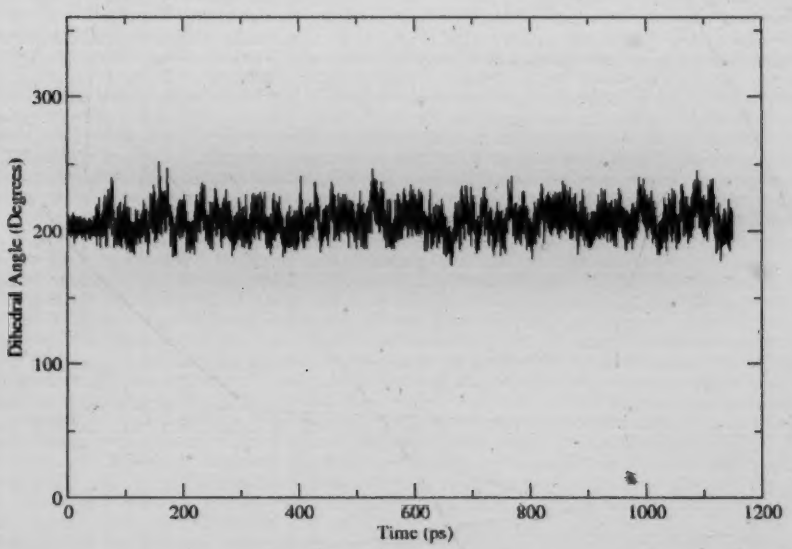

Figure 3.12: The main dihedral angle during 1 ns QM/MD simulation of VLB_1.

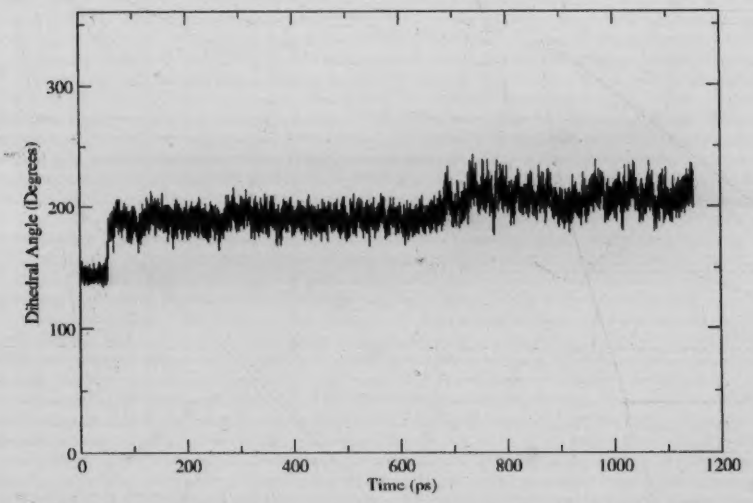

Figure 3.13: The main dihedral angle during 1 ns QM/MD simulation of VLB_2.

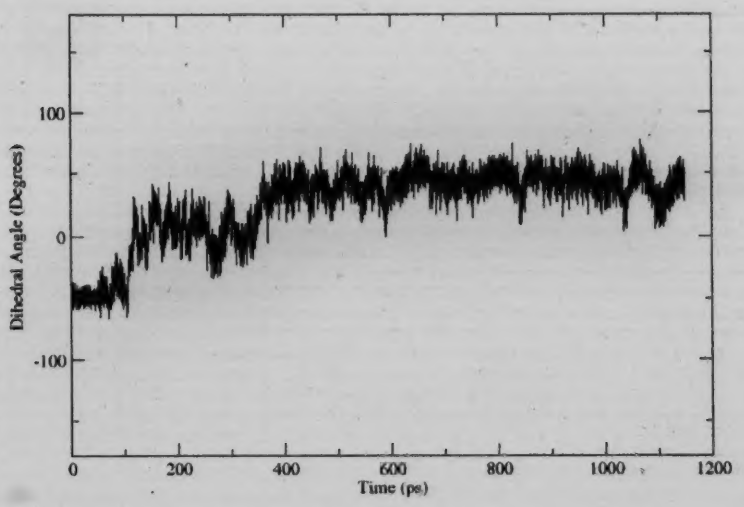

Figure 3.14: The main dihedral angle during 1 ns QM/MD simulation of VLB_3.

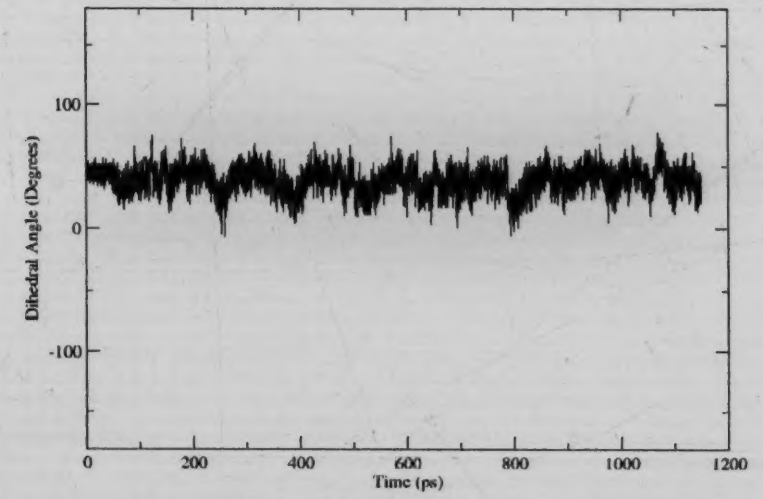

Figure 3.15: The main dihedral angle during 1 ns QM/MD simulation of VLB_4.

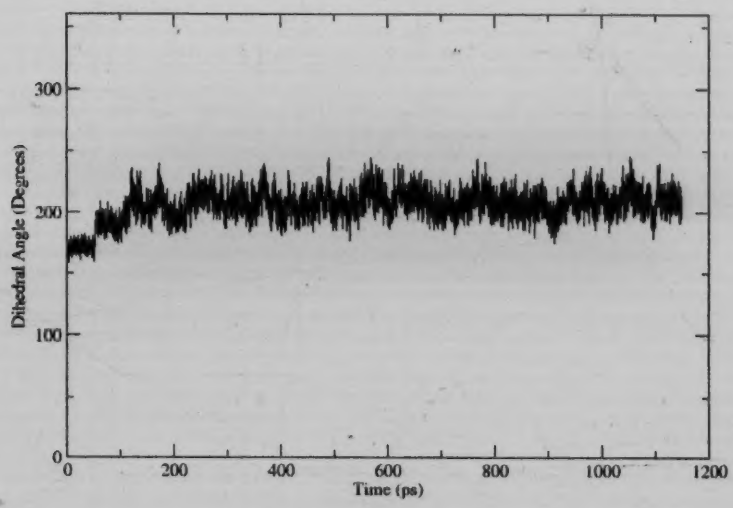

Figure 3.16: The main dihedral angle during 1 ns QM/MD simulation of VLB_5.

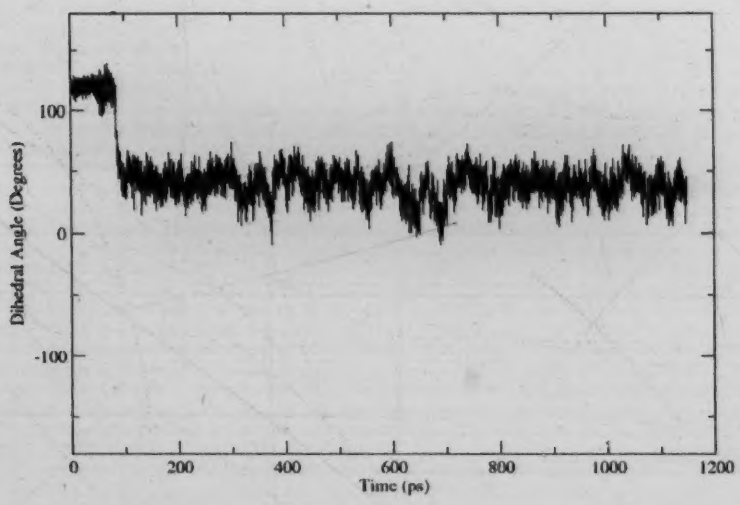

Figure 3.17: The main dihedral angle during 1 ns QM/MD simulation of VLB_6.

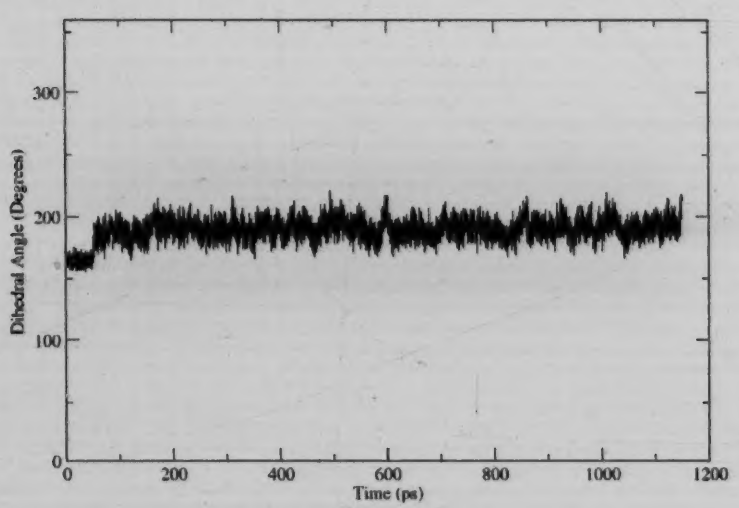

Figure 3.18: The main dihedral angle during 1 ns QM/MD simulation of VCR_1.

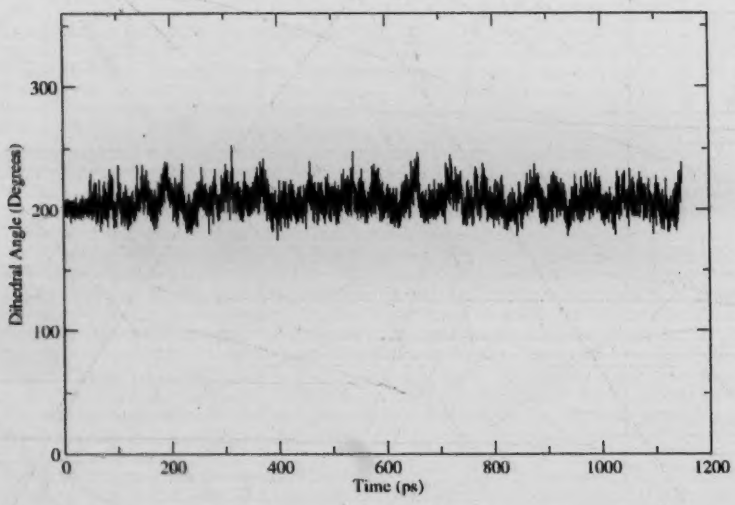

Figure 3.19: The main dihedral angle during 1 ns QM/MD simulation of VCR_2.

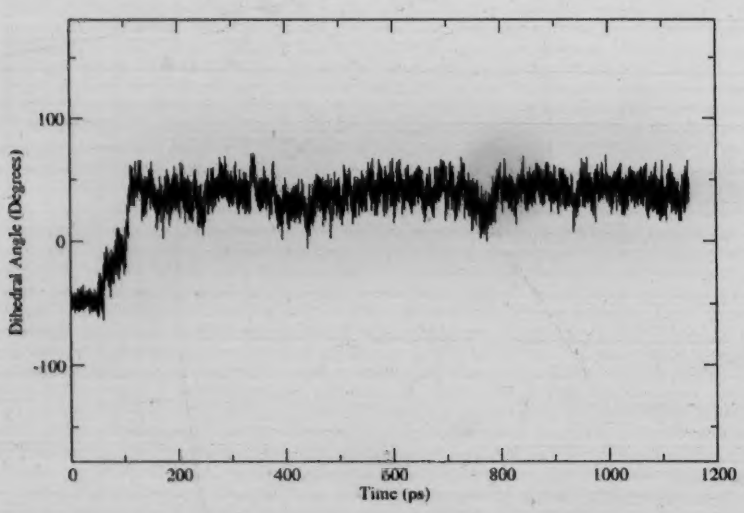

Figure 3.20: The main dihedral angle during 1 ns QM/MD simulation of VCR_3.

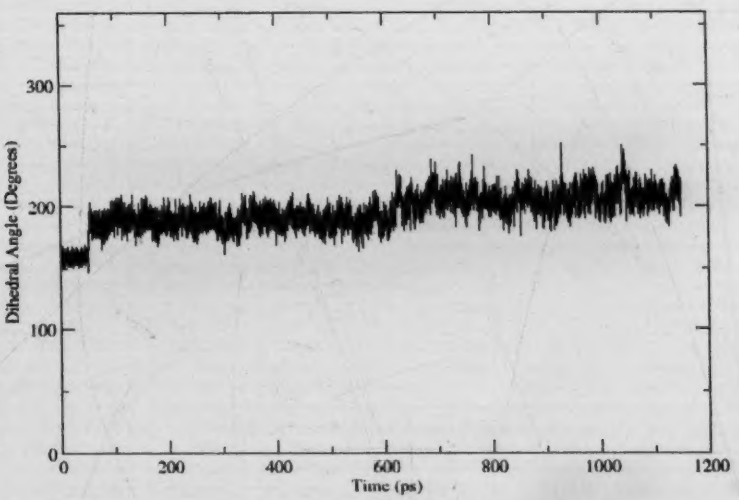

Figure 3.21: The main dihedral angle during 1 ns QM/MD simulation of VCR_4.

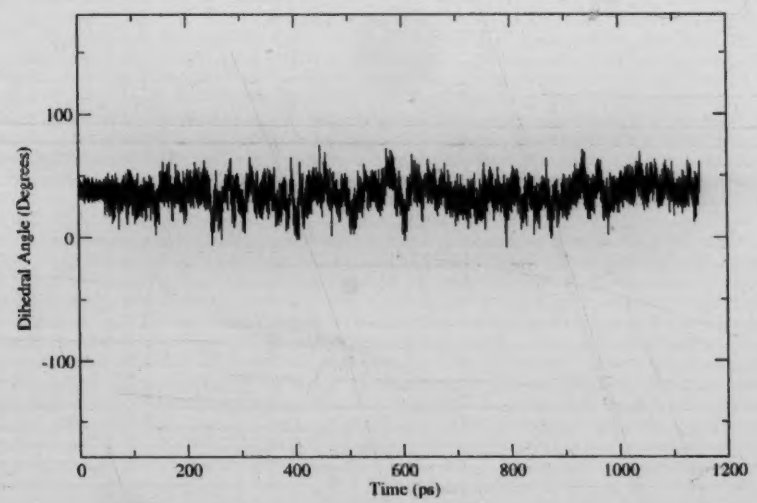

Figure 3.22: The main dihedral angle during 1 ns QM/MD simulation of VCR_5.

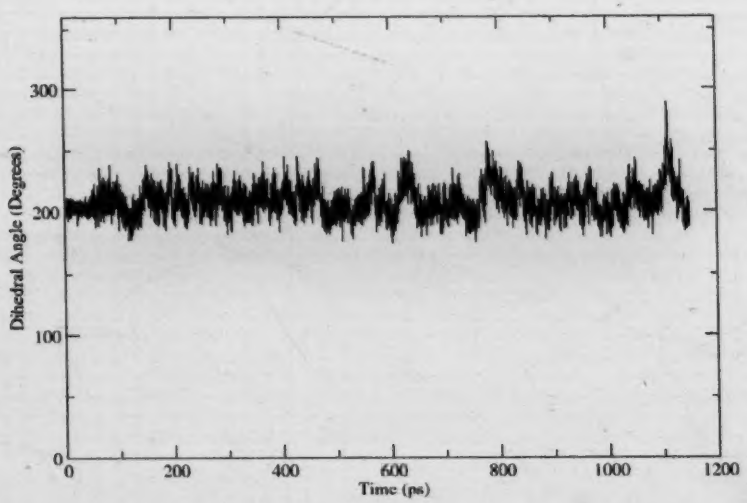

Figure 3.23: The main dihedral angle during 1 ns QM/MD simulation of VDE_1.

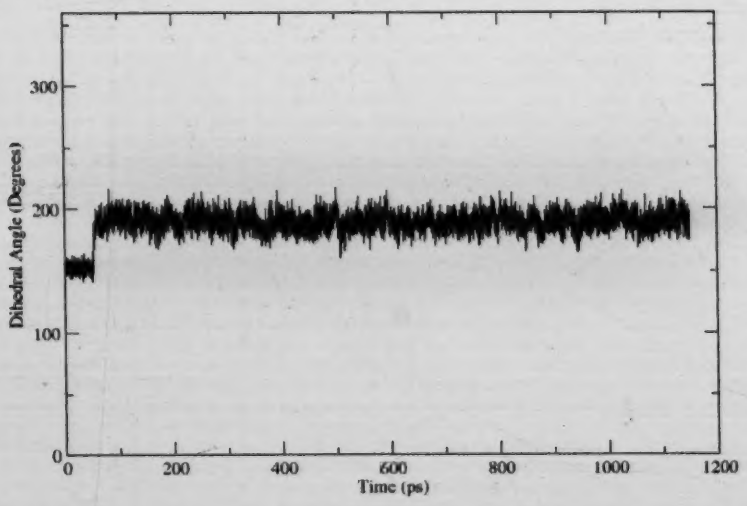

Figure 3.24: The main dihedral angle during 1 ns QM/MD simulation of VDE_2.

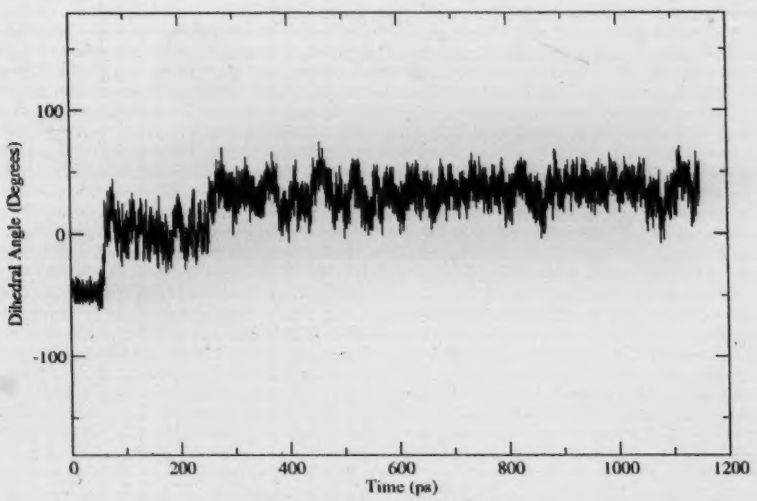

Figure 3.25: The main dihedral angle during 1 ns QM/MD simulation of VDE_3.

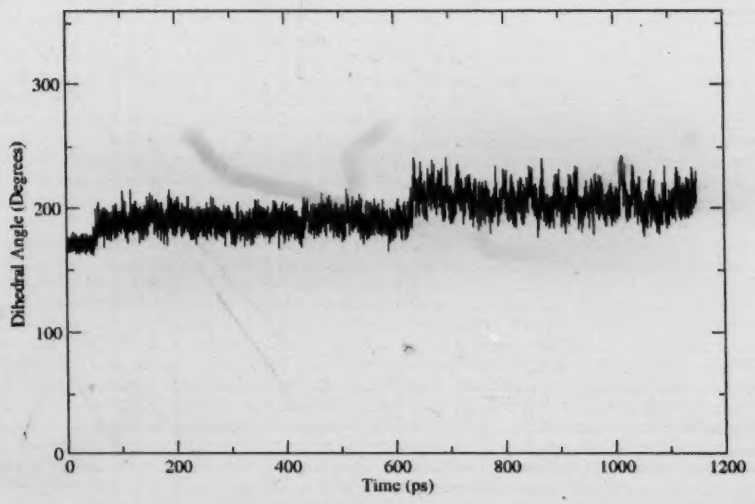

Figure 3.26: The main dihedral angle during 1 ns QM/MD simulation of VDE_4.

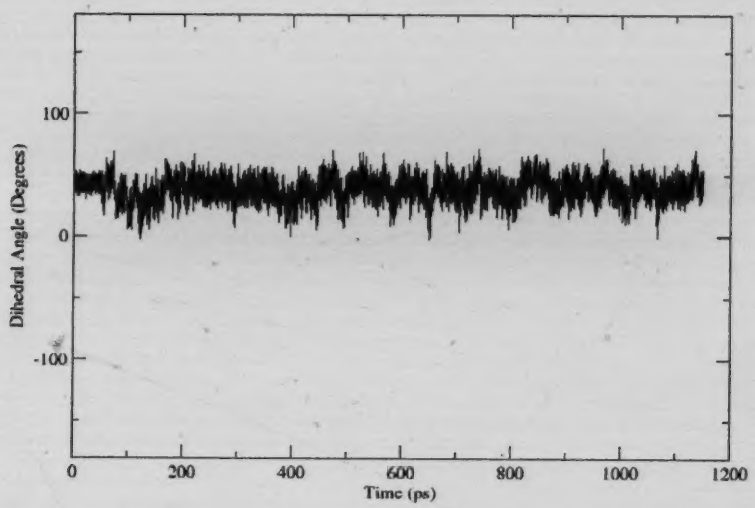

Figure 3.27: The main dihedral angle during 1 ns QM/MD simulation of VDE_5.

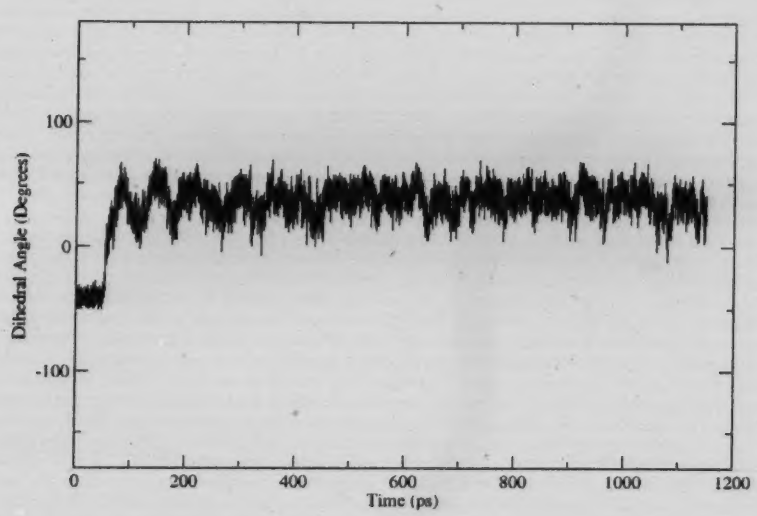

Figure 3.28: The main dihedral angle during 1 ns QM/MD simulation of VDE_6.

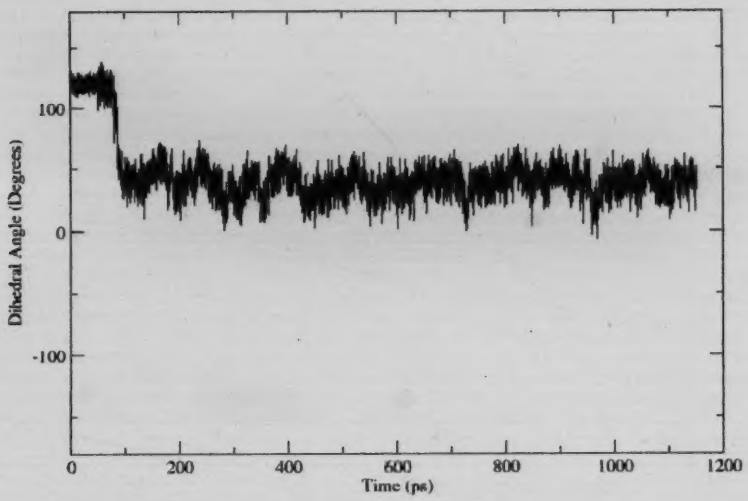

Figure 3.29: The main dihedral angle during 1 ns QM/MD simulation of VDE_7.

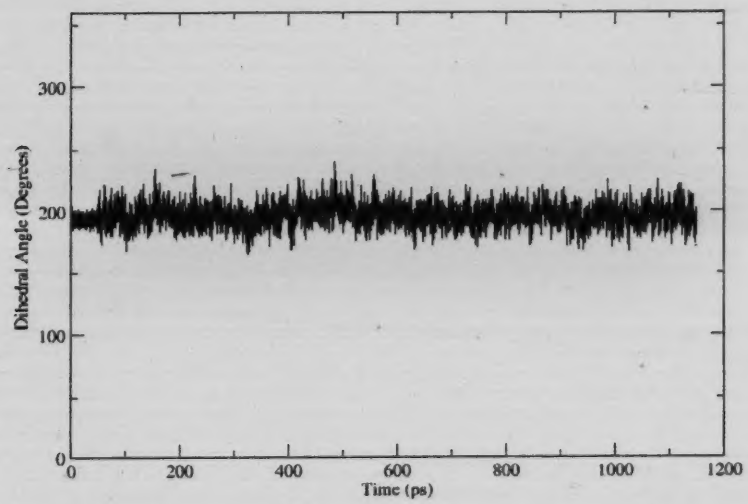

Figure 3.30: The main dihedral angle during 1 ns QM/MD simulation of VNO_1.

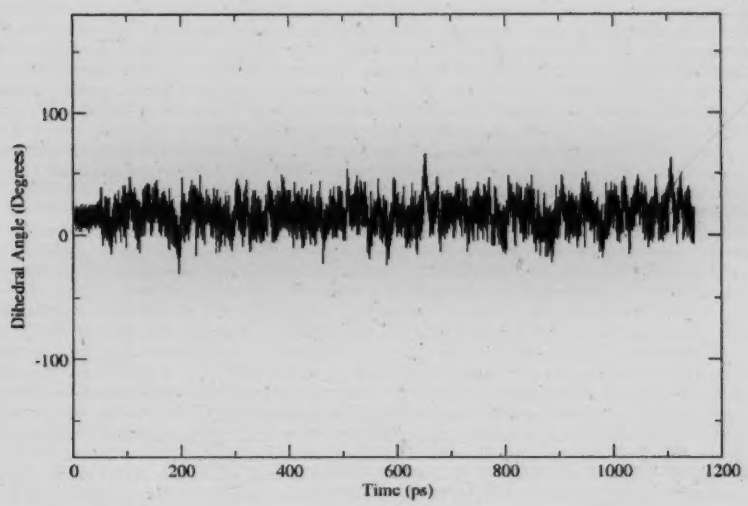

Figure 3.31: The main dihedral angle during 1 ns QM/MD simulation of VNO_2.

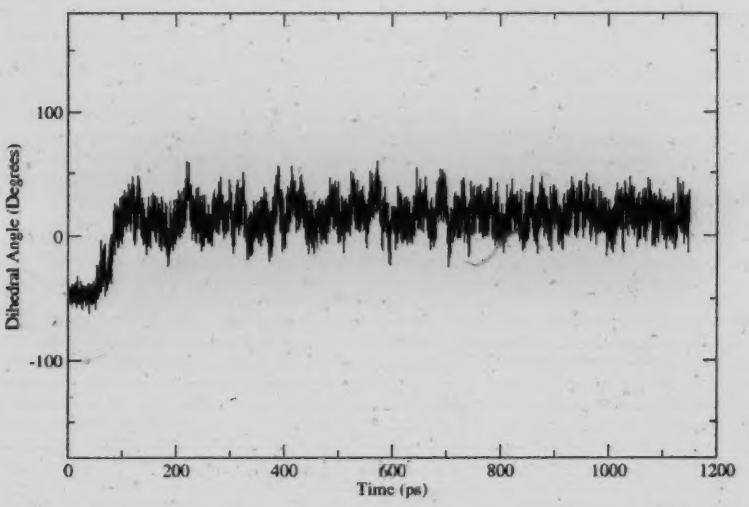

Figure 3.32: The main dihedral angle during 1 ns QM/MD simulation of VNO_3.

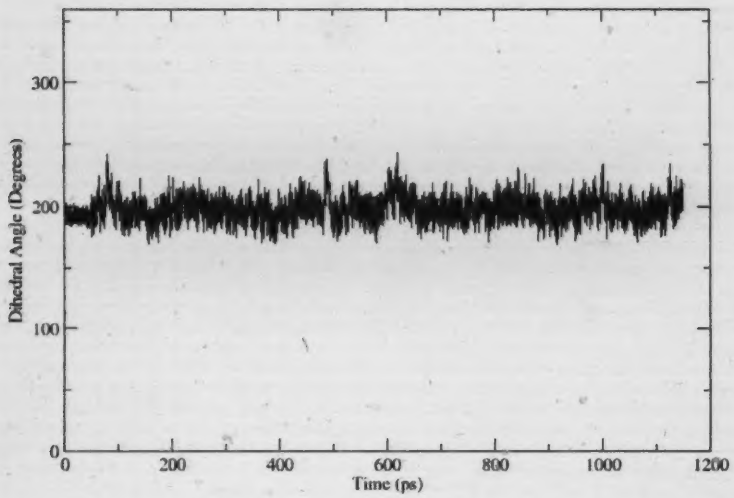

Figure 3.33: The main dihedral angle during 1 ns QM/MD simulation of VFL_1.

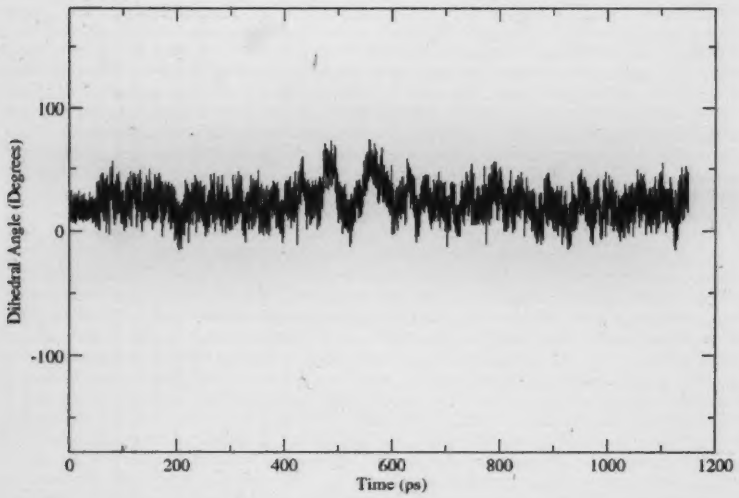

Figure 3.34: The main dihedral angle during 1 ns QM/MD simulation of VFL_2.

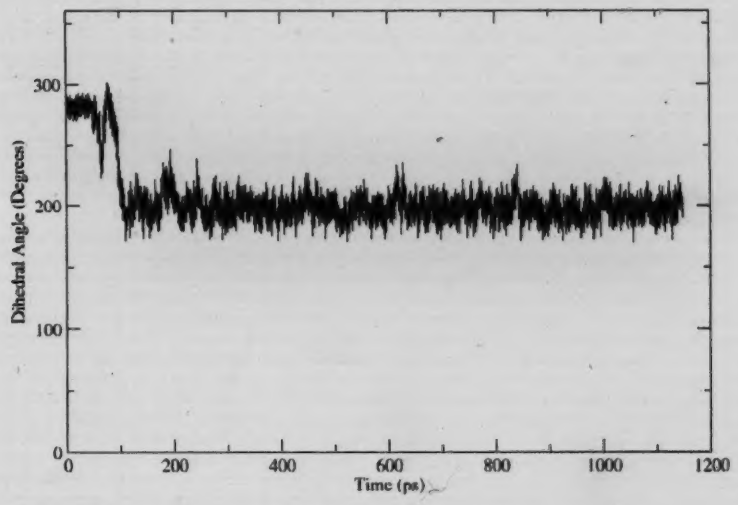

Figure 3.35: The main dihedral angle during 1 ns QM/MD simulation of VFL_3.

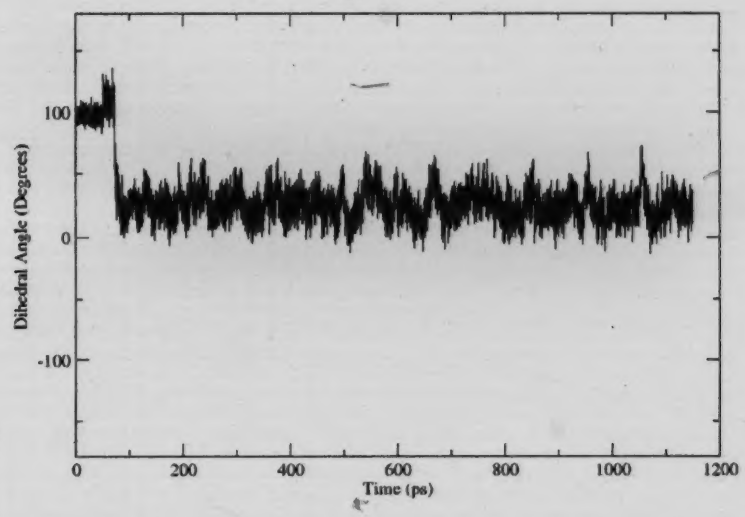

Figure 3.36: The main dihedral angle during 1 ns QM/MD simulation of VFL_4.

Table 3.3: The average properties of the *Vinca* alkaloids from 1 ns of QM/MD simulations. Error values, where available, are in parentheses behind each value. 'Angle' refers to the C17'-C18'-C15-C16 dihedral angle, and $r_{C4'-C10}$ is the average distance between atoms C4' and C10. R_G refers to the radius of gyration. E_{AM1} are AM1 energies are for the ligand only, and PE_{Tot} are total potential energies are for the entire QM/MM system, both in kcal mol^{-1}.

	Angle(°)	$\mathbf{r_{C4'-C10}}$ **(Å)**	$\mathbf{R_G}$ **(Å)**	$\mathbf{E_{AM1}}$	$\mathbf{PE_{Tot}}$
VLB_1	207.9 (0.8)	7.78 (0.04)	5.4 (0.005)	-192.6 (0.4)	-19096 (4)
VLB_4	38.8 (1)	9.64 (0.12)	5.5 (0.013)	-188.5 (0.4)	-19099 (3)
VCR_2	207.6 (0.8)	7.27 (0.10)	5.4 (0.008)	-226.5 (0.4)	-19124 (3)
VCR_5	36.4 (0.9)	9.03 (0.08)	5.5 (0.02)	-222.7 (0.5)	-19122 (3)
VDE_1	208.8 (1.3)	7.93 (0.03)	5.2 (0.004)	-120.6 (0.4)	-19011 (3)
VDE_5	39.4 (0.8)	9.66 (0.06)	5.3 (0.006)	-117.8 (0.4)	-19015 (4)
VNO_1	197.0 (0.7)	8.07 (0.04)	5.4 (0.006)	-122.5 (0.4)	-19023 (3)
VNO_2	16.6 (0.7)	9.22 (0.05)	5.4 (0.01)	-119.2 (0.4)	-19020 (3)
VFL_1	198.1 (0.7)	8.30 (0.04)	5.4 (0.004)	-243.3 (0.4)	-19140 (3)
VFL_2	23.5 (1.1)	9.42 (0.08)	5.4 (0.009)	-239.2 (0.4)	-19142 (4)

Chapter 4

Vinca Alkaloid-Tubulin Complexes

4.1 QM/MD simulations

With the structural and dynamical information on free *Vinca* alkaloids provided by the AM1 calculations and QM/MD simulations in Chapter 3, a comparison of the free drugs to the tubulin-bound drugs should provide valuable insight into the behavior and properties of the *Vinca* alkaloids. This chapter discusses the analogous simulations as performed in Chapter 3 with the *Vinca* alkaloids bound to a beta tubulin unit. QM/MD simulations were performed with the drugs described at the semi-empirical AM1 level.

4.1.1 Procedure

The 1Z2B PDB entry [2] was used as a starting structure for tubulin bound to a *Vinca* alkaloid drug. The PDB structure contains two whole tubulin heterodimers, Vinblastine bound to the *Vinca* domain, and an RB3 protein stathmin-like domain acting as a stabilizing structure. To simplify the model and reduce computational cost the alpha unit from the first heterodimer and the entire second heterodimer, as well as the stathmin-like domain, were deleted from the structure. What was left was a beta tubulin unit, a molecule of vinblastine in its bound conformation, and the nearby GDP molecule. This configuration is a basic model of the exposed +end of a microtubule, which is where *Vinca* drug binding is the most active and important. In order to most accurately reflect the conditions of a growing microtubule tip, the GDP molecule was converted, by the TELEAP program of the AMBER suite, to GTP by the addition of an extra phosphate group. The conversion was done according to a minimized template structure of free

GTP, and did not affect the existing GDP atoms. Similarly, minimized templates of the *Vinca* alkaloid drugs VCR, VDE, VNO, and VFL were prepared and used to mutate the VLB unit. In an attempt to limit human bias, no manual adjustments were made to the drug conformations. Thus, five structures were prepared: one for each drug, each with GTP at the tip.

The simulation procedure used for the tubulin-drug complexes is nearly identical to that used for the free drugs discussed in Chapter 3. Only the differences will be discussed here. The tubulin protein was described by the AMBER ff03 force field [90], and the GTP molecule was described by the force field parameters determined by Meagher [106].

The system was solvated with approximately 16000 TIP3P water molecules using TELEAP's `solvatebox` command. Randomly chosen water molecules were deleted and in their places 20 Cl^- and 39 K^+ ions were added to bring the system to a net neutral charge and an ion concentration of approximately 100 mM KCl – this will approximate the ionic conditions inside a typical human cell. The remaining water molecules were deleted. The solute and ion system was then resolvated with exactly 16000 water molecules. All other simulation details and parameters were the same as in the drug-only simulation.

The system was minimized first with all solute protein and drug atoms restrained: the drug was described with the same GAFF AM1-bcc parameters as in the free drug simulations. The second minimization stage was performed with most of the system unrestrained, except the drug was restrained again in an attempt to keep the conformation from changing through interaction with the protein. The warming of the solvent was then performed with protein and drug restrained. A solvent equilibration was then performed for a total of 250000 timesteps which gives a total of 500 ps. A full unrestrained equilibration, with the drug described at the AM1 level, was performed for a total of 0.1 ns. The production simulation was performed for 1 ns.

4.1.2 Results

Plots of the main dihedral angles varying versus time are presented in Figs. (4.1) to (4.5). The first 100 ps of each plot are from the end of the solvent equilibration stage where the bound drug is restrained. The next 100 ps are the unrestrained QM/MD equilibration, and the remaining 1 ns is the production simulation from which data was gathered. There were no major shifts in main dihedral angle, indicating that the starting

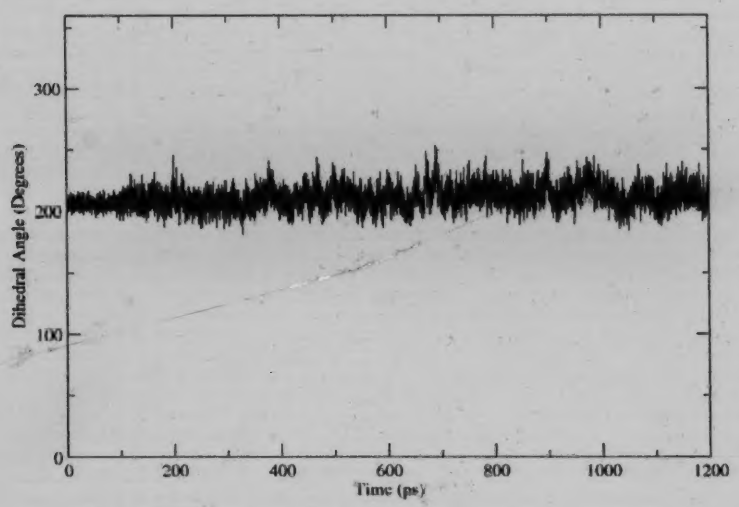

Figure 4.1: The main dihedral angle of VLB bound to beta tubulin over 1 ns.

structures were essentially stable to begin with. The averaged results of these simulations are presented in Table (4.1) along with the C4'-C10 distances. Also in Table (4.1) are averaged radii of gyration, AM1 energies of the solute drug, and overall QM/MM energies of the entire drug, protein, and solvent system. The blocking method was, as before, applied to calculate statistical errors. The method failed to converge when applied to the VLB main dihedral angle.

4.2 Functional Group Residue Interactions

A simple and direct quantification of drug-protein interactions is the analysis of contact between drug atoms and protein atoms. In particular, interactions between drug functional groups and amino acid side-chains are important as understanding these interactions can lead to new functional group derivative ideas in rational drug design or to understanding the effects of protein sequence mutations.

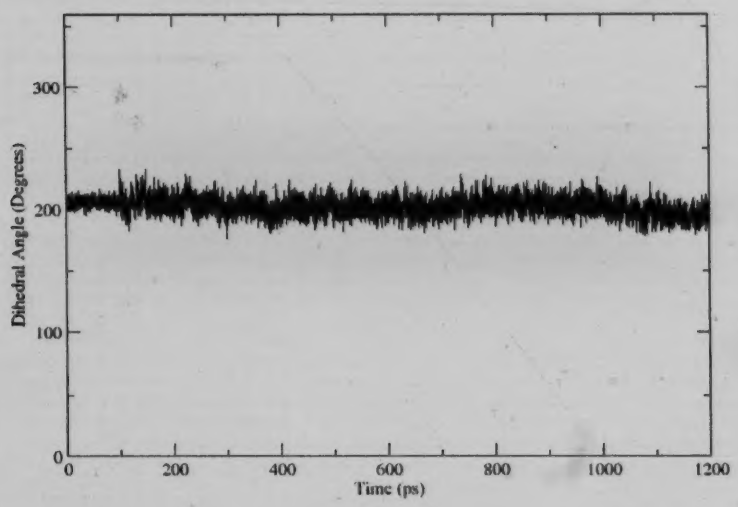

Figure 4.2: The main dihedral angle of VCR bound to beta tubulin over 1 ns.

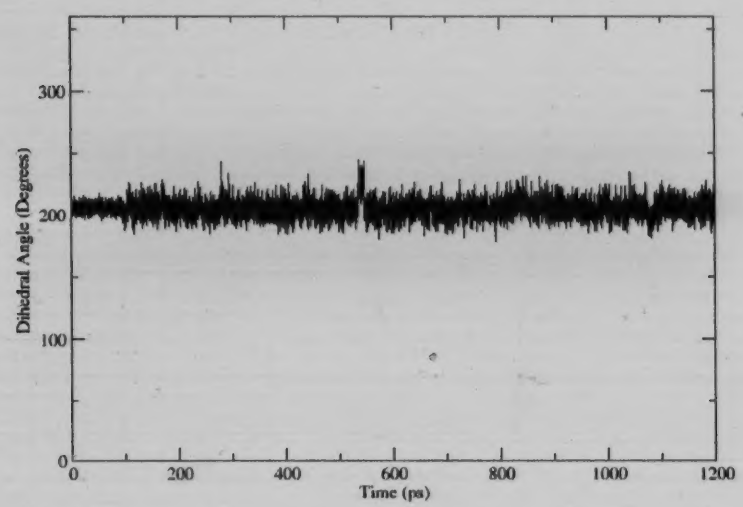

Figure 4.3: The main dihedral angle of VDE bound to beta tubulin over 1 ns.

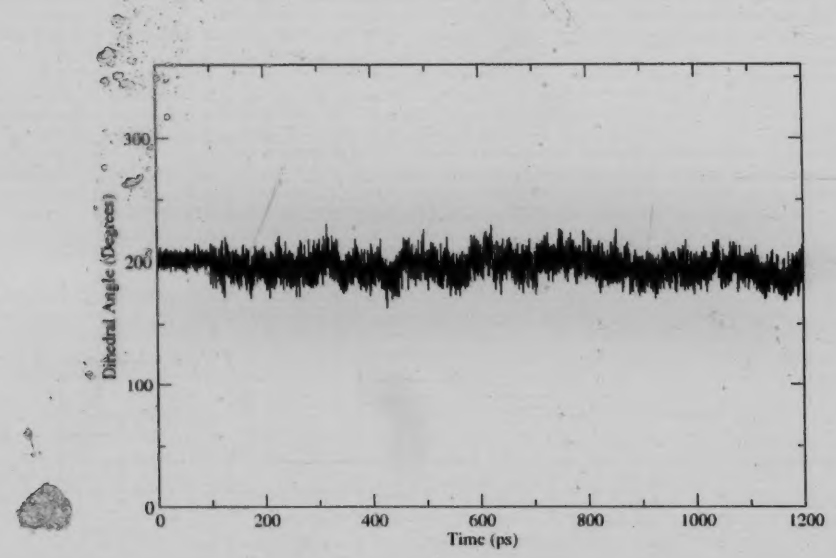

Figure 4.4: The main dihedral angle of VNO bound to beta tubulin over 1 ns.

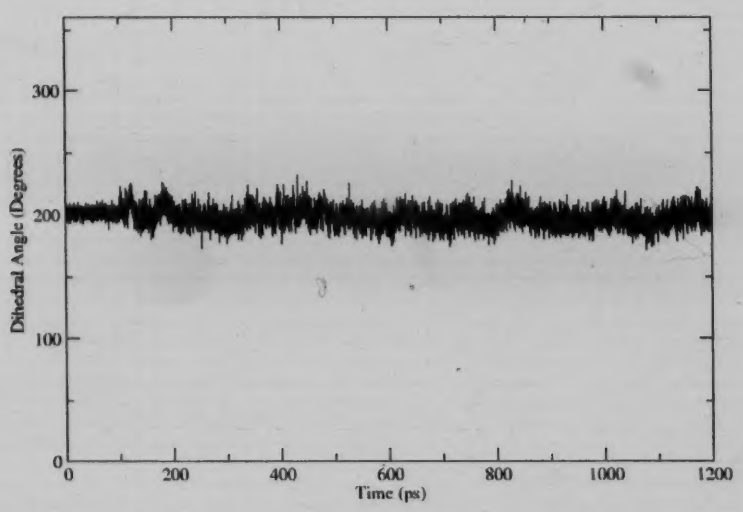

Figure 4.5: The main dihedral angle of VFL bound to beta tubulin over 1 ns.

Table 4.1: The average properties of the *Vinca* alkaloids bound to beta tubulin from 1 ns of QM/MD simulations. Error values, where available, are in parentheses behind each value. 'Angle' refers to the C17'-C18'-C15-C16 dihedral angle, and $r_{C4'-C10}$ is the average distance between atoms C4' and C10. R_G refers to the radius of gyration. E_{AM1} are AM1 energies for the ligand only, and PE_{Tot} are total energies for the entire QM/MM system, both in kcal mol^{-1}.

	Angle(°)	$r_{C4'-C10}$ **(Å)**	R_G **(Å)**	E_{AM1}	PE_{Tot}
VLB	211.4	7.63 (0.05)	5.3 (0.014)	-194.2 (1.2)	-166531 (40)
VCR	201.6 (0.7)	8.29 (0.05)	5.4 (0.006)	-241.0 (0.8)	-166605 (30)
VDE	205.1 (0.4)	7.70 (0.03)	5.1 (0.003)	-131.9 (0.5)	-166471 (30)
VNO	195.6 (0.9)	7.94 (0.04)	5.4 (0.003)	-136.7 (0.6)	-166452 (20)
VFL	197.4 (0.6)	8.12 (0.06)	5.4 (0.011)	-247.2 (0.5)	-166575 (16)

4.2.1 Procedure

Coordinate snapshots were taken 10000 times (once every 0.1 ps) during the production simulation. The `checkoverlap` command in the PTRAJ utility was used to record all pairs consisting of a drug atom and a protein atom that were within 4 Å of each other. These results were then summed over time and any interaction that was present for more than 5000 snapshots, or 50% of the 10 ns simulation, in total were considered significant. This analysis was done on a Apple PowerMac G5 Quad 1.9 GHz computer. These results are presented in Table 4.2 below.

4.2.2 Results

Table 4.2: The atom-pair interactions between bound drug and tubulin. Numbers entered are residence populations as a percentage of the total simulation. Please refer to Fig. 1.6 and the footnotes here for drug atom names. Protein atom names are from the standard AMBER FF03 force field. A dash indicates that the interaction was not significant and N/A indicates that the particular atom is not present in that drug.

	VLB	VCR	VDE	VNO	VFL
GLN11(CB)-C21'	65	-	-	-	-
PRO173(C)-C10	-	-	-	50	-
PRO173(O)-C10	-	-	63	71	-
PRO173(O)-C11	-	70	82	-	-
LYS174(C)-C11	-	70	66	-	73
LYS174(CA)-O34 [1]	N/A	55	N/A	N/A	N/A
LYS174(CD)-O34	N/A	61	N/A	N/A	N/A
LYS174(O)-C11	51	94	81	-	89
LYS174(O)-C22	51	-	51	-	-
LYS174(O)-C4'	-	-	-	61	-
LYS174(O)-C5'	-	-	-	56	-
LYS174(O)-C19'	-	-	-	60	-
VAL175(C)-C11	-	55	-	-	-

continued on next page

[1] O34 is only present in VCR, it is the oxygen atom of the N1-formyl group

Table 4.2: *continued*

	VLB	VCR	VDE	VNO	VFL
VAL175(N)-C11	-	63	-	-	51
VAL175(O)-C10	-	68	-	-	68
VAL175(O)-C11	-	60	-	-	-
ASP177(CB)-C8	-	57	-	-	-
ASP177(OD2)-C8	-	-	69	-	-
TYR208(CD2)-C3'	54	55	-	-	69
TYR208(CE2)-C3'	52	-	-	65	-
TYR208(OH)-C16	-	-	-	57	-
TYR208(OH)-C17	83	88	89	75	62
TYR208(OH)-C22	69	75	84	-	-
TYR208(OH)-C33[2]	76	71	83	72	90
CYS211(CB)-C21'	65	-	-	-	-
THR218(O)-C26'[3]	-	-	-	-	65
PRO220(C)-C21'	-	-	53	-	-
PRO220(C)-O22'	-	85	52	N/A	N/A
PRO220(CD)-C26'	-	-	-	-	51
PRO220(CD)-O25'[4]	-	-	-	-	55
PRO220(CG)-C33	-	70	-	-	-
PRO220(CG)-O32[5]	-	58	-	-	-
PRO220(O)-C1'	-	56	-	-	-
PRO220(O)-C4'	-	58	-	-	57
PRO220(O)-C7'	-	-	-	-	73
PRO220(O)-C21'	-	-	51	-	-
PRO220(O)-O22'	52	100	98	N/A	N/A
PRO220(O)-O25'	-	64	-	-	-
THR221(C)-C5'	53	-	-	-	-
THR221(C)-C20'	-	-	60	-	-
THR221(C)-C21'	-	-	89	58	-

continued on next page

[2]C33 is the C18' formyl carbon
[3]C26' is in the alcohol portion of the C18' ester group
[4]O25' is in the alcohol portion of the C18' ester group
[5]O32 is the C18' formyl oxygen

Table 4.2: *continued*

	VLB	VCR	VDE	VNO	VFL
THR221(C)-F2 [6]	N/A	N/A	N/A	N/A	57
THR221(CA)-C21'	-	-	60	-	-
THR221(CA)-O22'	-	78	50	N/A	N/A
THR221(N)-C21'	-	-	61	-	-
THR221(O)-C21'	-	-	74	-	-
THR221(O)-F2	N/A	N/A	N/A	N/A	55
TYR222(CE1)-C20'	-	76	-	-	-
TYR222(CE2)-C5'	-	77	85	-	-
TYR222(CE2)-C20'	-	62	-	-	-
TYR222(CZ)-C5'	-	51	78	-	-
TYR222(CZ)-C20'	-	81	-	-	-
TYR222(N)-C7'	56	-	-	-	-
TYR222(N)-C20'	-	-	74	-	-
TYR222(N)-C21'	-	-	59	-	-
TYR222(OH)-C5'	-	-	58	66	-
LEU225(CD1)-C21'	-	-	69	-	-
LEU225(CD1)-F1[7]	N/A	N/A	N/A	N/A	54
LEU225(CD1)-F2	N/A	N/A	N/A	N/A	65

[6] F2 is present only in VFL. It is attached to the C4' ethyl group.
[7] F1 is present only in VFL. It is attached to the C4' ethyl group.

Chapter 5

Discussion

5.1 Free Form *Vinca* alkaloids

5.1.1 Potential Energy Surface

Each *Vinca* alkaloid has a similar projection of its potential energy surface onto the main dihedral angle. VLB, VCR, and VDE each have a primary minimum at an angle of just over 200°, and a secondary minimum near 37°. VNO and VFL have similar energy profiles, but have primary and secondary minima at lower angles near 195° and between 10-20°. Structural analysis shows that the top portion of each drug undergoes distortion due to steric crowding in the high energy conformations. This is expected as the top portion has a flexible 8- or 9- membered ring, and that is exactly where the distortion occurs. The bottom half of the drug is much more rigid, due to its polycyclic nature containing smaller rings. The two groupings of conformations mentioned above likely occur due to similarity in the top portions of the drugs: VLB, VCR, and VDE each have the same top structure, but VNO and VFL both have a smaller 8-membered ring instead of a 9-membered ring. It was theorized by Hunter [19] that the conformation of VLB would not be dependent on side-chain interaction, and this theory appears to be confirmed here: it is distortion of the top portion that gives rise to conformational preference in the *Vinca* alkaloids.

The surfaces presented here were not smooth, as would be expected in a torsional energy profile, but rather jagged with many spikes. The most reasonable explanation is that the spikes are due to the problem of the system becoming caught in a local minimum and not a global minimum (for the torsional restraint applied), which translates to a falsely high energy for that torsional angle being displayed. The goal of calculating

these torsional energy profiles was to provide a guide for the conformation of the *Vinca* alkaloids. Reasonable effort was taken to smooth the profiles as much as possible. While the energy spikes are not ideal, the results here are sufficiently smooth to satisfy the initial goal: it is clear that one conformation near 200° is strongly preferred. A secondary minimum in the range of 10-40° exists, but is of higher energy and is largely insignificant (see the Section 5.1.3). The observation of this secondary minimum is consistent with the results of [22], where computational modeling found a minimum of VLB at 39°, which was discounted based on experimental data in favour of a minimum at 207°.

5.1.2 MD simulations

The 0 K potential energy surfaces predict only two stable conformers for each *Vinca* alkaloid. The QM/MD simulations performed follow those predictions very closely: regardless of the starting conformation the structure collapses to one of the two minima. The average angles in the MD simulations are within a few degrees of the minima in the 0 K energy surfaces, indicating that solvation and non-zero temperature do not considerably affect the structures predicted by the 0 K minimizations of the *Vinca* alkaloids. Thus, the grouping mentioned above (VLB, VCR, VDE, vs. VNO, VFL) still holds in aqueous MD simulations. The average C4'-C10 distances are all within 1.1 Å of each other, thus the interaction between each drug and the *Vinca* domain is expected to be very similar in each drug as the relative orientation of the tubulin-interacting regions of the top and bottom portions are comparable. The radius of gyration for each drug is approximately 5 Å, and this may be relevant for future comparison to certain experiments such as electron scattering measurements. The energy of the QM regions, which are the AM1 energies of the drugs alone, were calculated to a low enough error for comparison to be significant. However, the total potential energies of the system were not as accurately determined and their differences were smaller than the errors of each measurement. Longer simulation times could potentially lower the errors.

The main dihedral angle is the most important geometric parameter of each of the *Vinca* alkaloids. When comparing the conformers of one drug to each other, the main dihedral is the best way to classify the conformers. However, when making comparisons between the drugs, the main dihedral angle may not be the best parameter as the angle depends on the structures of the top and bottom portions and these may be different, thus making a main dihedral comparison ill-defined. Furthermore, the global effects of a change in the main dihedral angle may not be the same in each drug due to the aforementioned structural differences. For these reasons, the additional parameter of the

C4'-C10 distance was used to aid in comparison between drugs. The equilibrium angles differ by a range of about 10° and this results in a C4'-C10 distance range of about 1.1 Å, which may cause differential binding properties, but is far from a guarantee to do so.

5.1.3 Conformation Populations

Based on the torsional potential energy projections presented in Figs. 3.1 to 3.5, it is evident that of the two minimum angle conformations the conformation at approximately 200° is the preferred conformer as its energy is lower. However, the relative populations of the two conformers for each drug will be informative. Assuming Boltzmann distribution of states, the relative probability between the two is:

$$\frac{N_1}{N_2} = e^{\frac{E_2 - E_1}{k_B T}} \tag{5.1}$$

where N_1 and N_2 are the populations of the primary and secondary conformers, respectively, E_1 and E_2 are their energies, k_B is the Boltzmann constant, and T is the absolute temperature. If the *in vacuo* 0 K potential energies are used, for each drug except for VCR, the 200° conformation will be present almost exclusively at 310 K. Using the 0 K potential energy difference, the secondary conformation of VCR will be populated by approximately 20% of the VCR molecules in aqueous solution. However, the averaged AM1 energies from the QM/MD simulations of the *Vinca* alkaloids in TIP3P water must be considered. The energy differences between conformers in the MD simulations are approximately the same as in the 0 K energy surfaces, except for VCR, where the secondary conformer is much higher in energy than the primary. The AM1 simulations then predict that each *Vinca* alkaloid will be present almost exclusively in the primary conformation near a main dihedral angle of 200°. It would be reasonable to conclude that the 0 K potential energy surface is less trustworthy than the MD simulations as 0 K energy minimizations of complicated molecules such as the *Vinca* alkaloids are prone to finding local minina; but, due to the finite temperature, and hence dynamical nature, of the MD simulations the problem of local minima is less likely to occur. The primary minimum energy structure of VCR is then most likely not a true minimum. Another possible explanation for this difference is solvation: the 0 K potential energy calculations were done *in vacuo*. VCR differs from the other drugs in the conversion of an apolar methyl group to a polar formyl group, thus electrostatic screening by solvent or hydrogen bonding may explain the difference. Visual inspection reveals a hydrogen bond from the formyl O to the C3 hydroxyl group in both of the 0 K minima. This hydrogen bond is not present in the QM/MD simulation of VCR. The aqueous, finite temperature simulation

is a more realistic and relevant representation of the *Vinca* alkaloids than an *in vacuo*, static, model. Therefore, the remaining discussion will focus primarily on the primary minimum structure found. However, the torsional potential energy surfaces are still valid as a guide to discuss the *Vinca* alkaloid conformation.

5.1.4 *in vacuo* Dipole Moments

The primary minimum structures all have dipole moments of about 2 D, with the exception of VCR which has a dipole moment of 3.6 D. This difference may be due to the previously discussed inaccuracy in the primary minimum of VCR. Even if the value for VCR were accurate, each of the drugs may be described as polar, but not very polar. This is shown in the fact that the drugs are of fairly low water solubility unless in acidic conditions. The similar, and moderate, dipole moments indicate that relevant properties of the drugs, of particular interest being cell membrane transport, should be analogous.

Of interest is how VFL is not significantly more polar than the other drugs, despite the presence of two very electronegative fluorine atoms. This is explained by examining the availability of electron density near the fluorine atoms. The fluorine atoms are bound to an ethyl group that is attached to a singly bonded organic ring. The absence of lone pairs, multiple bonds, and conjugation here means that there is little free electron density to move towards an electronegative region of the molecule; despite being highly electronegative, the fluorine atoms have no free electron density to attract. Even though the fluorine atoms of VFL do not affect the global electrostatic properties of the molecule, this is not to say that the flourine atoms will not affect the local properties. In particular, the interaction with tubulin may be altered, especially in light of the evidence showing that this region is buried when bound to tubulin: Rai's study using Ant-VLB and Gigant's crystal structure both show this.

5.2 The *Vinca* Alkaloids bound to Beta Tubulin

5.2.1 MD Simulations

The 1Z2B crystal structure [2] used as starting coordinates for the MD simulations of all five *Vinca* alkaloids contains VLB in a conformation (main dihedral is 213°) similar to the equilibrium conformers found in the previously discussed MD simulations. This is a good indication that the 1Z2B structure is a reasonable starting point for studies of the *Vinca* alkaloids bound to tubulin. The 1Z2B structure shows a binding mode

consistent with the fluorescence studies of Rai [17] and Chatterjee [37]: the the C4' atom is buried and the C4 functional group is solvent exposed. The main dihedral angles in the simulations are all near their free form values, and cover a slightly broader range of about 16°. The blocking method failed when applied to the main dihedral average of VLB; the other dihedral simulations had reasonably low values. The C4'-C10 distances are also similar, but their range is smaller than for the free drugs: only 0.7 Å. Visual inspection of MD trajectories shows that the important features of the Bau and Jin crystal structure [1] including the chair conformation of the piperidine ring and near co-planarity of the indole and 8- or 9-membered rings in the top portion are present. The radii of gyration were also all comparable to the free form values. The AM1 energies were calculated to a low enough error to be significant. Further discussion will be found in the next section.

5.2.2 Drug-protein pair-wise interactions

Table 4.2 presents the most common pair-wise atomic interactions between the drug and protein in the bound form MD simulations. This study is by no means an exhaustive examination of the interactions present, but it was intended to provide insight towards how the *Vinca* alkaloids interact with beta tubulin and how the binding differs between the drugs based on the differential molecular structures.

Several interactions appear in most, if not all, drug simulations. The C11 methylene atom interacts with the carbonyl O of the backbone of LYS174 in all drugs except VNO – this is expected to be an unfavorable apolar-polar interaction. The side-chain hydroxyl group of TYR208 interacts with both the benzene C17 atom and C33 (of the C16 formyl group) in all drugs – this is expected to be favorable. Also of note is the interaction between the TYR208 hydroxyl and C22 (attached to the N1 atom) in VLB, VCR (in a formyl group), and VDE. The TYR208-C22 interaction is interestingly absent in VNO and VFL, despite that these drugs are identical in the bottom portion to VLB. Similarly, the hydroxyl O22' interactions with the PRO220 backbone oxygen are present in VLB, VCR, and VDE, but this atom is absent in VNO and VFL. Also, O22' interacts with the C-alpha of THR221 in VCR and VDE in a potentially unfavorable interaction.

In addition to TYR208, C22 also interacts with LYS174 in VLB and VDE. The C22 atom is of relevance due to the suggestion made by Magnus [53] that VLB, and not VCR, can act as a methylating agent of cysteine residues through C22. The only cysteine residue known to be in the *Vinca* domain is CYS211, near TYR208. While

an interaction between C22 and CYS211 is not demonstrated here, it remains possible that a rearrangement of VLB in the binding domain may initiate such an interaction, thereby making methylation possible. However, since CYS211 is shown to interact with the C21' atom of an ethyl group which is buried deep inside the protein, a C22-CYS211 interaction is not likely, and Magnus's theory may not be applicable.

The O34 atom is unique to VCR and is in the formyl group attached to N1. This atom appears to be of at least moderate importance through its interactions with LYS174. The conversion of the C22 methyl group in the other drugs to a formyl group in VCR may not affect the C22 interactions in VCR, though, as is seen in the TYR208 hydroxyl interactions with C22 in VLB, VCR, and VDE. The two ester groups at C3 and C4 in VLB undergo major reductions in the conversion to VDE. However, none of the deleted or added atoms were involved in any significant interactions – this region of the *Vinca* alkaloids does not appear to interact with beta tubulin. Also, the C7' and C8' atoms of the 8- or 9-membered ring in the top portion do not interact directly with beta tubulin, making the ring shortening in VNO and VFL of no consequence in this fashion. This ring shortening may have direct consequence through conformational effects, though. The C3'-C4' double bond in VNO may have some effect on binding compared to the singly bonded analogues in the other drugs, as is evidenced in the interactions with LYS174, TYR208, and PRO220. The most exotic conversion in the five *Vinca* alkaloids studied here is the addition of 2 fluorine atoms on the C4' ethyl group of VFL. This region of the drug is very important in the binding mode studied here. As might be expected, these atoms do directly participate in drug-protein interactions. The backbone of THR221 and side-chain of LEU225 are involved.

The *Vinca* domain has several sites for isotype mutations contained within it [29]. It is worth noting that only one significant interaction between a drug bound to beta tubulin was found to involve the mutation sites. This interaction is between VFL's C18' ester group and the backbone of THR218, which is where the βIII, βVII, and βVIII isotypes (using the notation of Huzil [29]) undergo divergent mutations. This interaction may be important in the action of VFL. However, it is an interaction with the backbone of the residue, not the side-chain, so side-chain mutations may not affect this interaction. Overall, the design of isotype specific *Vinca* alkaloids may require more subtle analysis than direct atom-atom interactions. While the interaction residues identified here are consistent with previous enumerations of the *Vinca* domain[1] the simulations performed

[1]found to be residues β173-β211 in [17] and β172-β177,β208-β225 in [29]

here suggest that residue β11 should also be included in the binding domain.

5.2.3 Comparison of Bound to Free Form

The main dihedral angles of the equilibrium free *Vinca* alkaloids are all very similar to the angles observed in the bound form. The largest difference is of 6.0° in VCR. With the exception of a 1 Å difference in VCR, the C4'-C10 distances are also very similar in free and bound forms, differing by at most about 0.2 Å. With the possible exception of minor rearrangements of functional groups, these measures indicate that the gross conformation of each *Vinca* alkaloid is the same bound as it is free. As no major reorientation of the *Vinca* alkaloids is needed in order to bind, this suggests a simple one- or two-step binding mechanism. A one step mechanism would simply be a collision of the *Vinca* alkaloid surface with the protein surface. A two step mechanism may involve the collision and binding of either the top or bottom portion of the drug followed by the interaction and binding of the other portion.

The AM1 energies are all lowered in the bound form compared to their equilibrium free form, VLB by only the small amount of about 2 kcal mol^{-1} and up to about 16 kcal mol^{-1} in the case of VNO. This is indicative of a stabilization of the *Vinca* alkaloids upon binding. Broader energy difference measurements, such as protein deformation energies, drug desolvation energies, or overall binding energies or enthalpies are not possible within the simulations performed. The AM1 energy decrease upon binding can be seen to indicate that binding is favorable. As no major conformational changes occur in the *Vinca* alkaloids upon binding, it can be hypothesized that the AM1 energy decreases are due to interaction with tubulin. This may be through the formation of favorable polar-polar (or apolar-apolar) functional group-side-chain interactions or through the removal of hydrophobic regions of the drug from contact with water.

Keeping the above points in mind, the binding event of each of the five *Vinca* alkaloids to beta tubulin will likely be a fast, simple event that is driven by non-covalent interactions. This agrees with the previously proposed mechanism of rapid and reversible binding [36]. This study did not find any reasons for the binding of VFL to the *Vinca* domain to be considerably different than the other four *Vinca* alkaloids studied, as was found by Kruczynski [47]. However, this is not to say that more in-depth simulations or calculations would not discover such reasons – this study is only a model of a single bound state and does not model the binding dynamics of the *Vinca* alkaloids.

5.2.4 Validity of Model

The 1Z2B crystal structure is the only set of high-resolution coordinates of a *Vinca* alkaloid bound to tubulin known. The binding domain and mode in this structure agree well with other studies on the binding of the *Vinca* alkaloids. The calculations and simulations on the free *Vinca* alkaloids demonstrate that only one conformation for each will be present in any significant quantity at 310 K. As this equilibrium conformation closely resembles the conformation found for VLB in 1Z2B, the binding mode in 1Z2B must be considered highly likely for all the *Vinca* alkaloids studied. Furthermore, in the MD simulations of the *Vinca* alkaloids bound to tubulin, the complexes do not destabilize in any apparent way or expel the *Vinca* alkaloid from the binding domain, indicating at least basic stability. Of course, longer simulation times would strengthen this last statement, but these were not possible within this project.

5.3 Conclusions

The *Vinca* alkaloids studied here have been demonstrated to have one preferred equilibrium and minimum energy conformation with main dihedral angle near 200°. It is then highly likely that this conformation is prerequisite to a *Vinca* alkaloid congener to function as an anti-mitotic agent. Modifications to the drug in rational drug design that disrupt this equilibrium conformation may then destroy the activity of the drug. Similarly, modifications that work within and strengthen this preferred conformation may have a better chance of efficacy. The structural feature that gives rise to this preferred conformation is the sterically induced distortion of the upper portion of the *Vinca* alkaloids. The unique equilibrium angles found in VNO and VFL are likely due to the modifications made to the upper portions of those drugs. Therefore, the upper portion backbone of the *Vinca* alkaloids is of particular importance as changes may seriously affect the activity of a *Vinca* alkaloid drug.

The MD simulations performed indicate that no major conformational changes occur in the *Vinca* alkaloids on binding to beta tubulin. This supports the previous hypothesis that *Vinca* alkaloid binding is rapid and reversible. Furthermore, the drug-protein interaction is shown to likely be through weak surfacial interactions, which also supports a rapid and reversible mechanism.

5.4 Future Work

While the semi-empirical AM1 model appears to accurately model the *Vinca* alkaloids, it is computationally intensive. Longer simulations, as accessible by a force field representation of the *Vinca* alkaloids, would allow for new simulations to be performed and new data to be gathered. The development and testing of a force-field which accurately represents the *Vinca* alkaloids would be a worthwhile next step in this research. Such development may involve the use of an existing general force-field, the *de novo* fitting of parameters, or a combination of both. Such a force field should reproduce the main dihedral angle properties observed.

An improvement to the AM1 0 K torsional potential energy surface would be a torsional free energy surface obtained from MD simulations of the *Vinca* alkaloids. A Free Energy of Perturbation style method, involving torsional restraints and thermodynamic integration may be used to accomplish this.

While the binding mode and drug conformation used here are both valid, this does not preclude the existence of different modes or bound conformations; these may arise from more complicated binding mechanisms than the simple, rapid one supported by this study. Such binding modes may be relevant depending on the drug concentration, environmental or solvent conditions, tubulin isotype, or especially if it is a new *Vinca* alkaloid congener being studied.

With the parametrization of an accurate force-field model of the *Vinca* alkaloids, a free energy of perturbation, or other, experiment could be performed to determine accurate free energies of binding for each *Vinca* alkaloid. This would provide direct comparison to binding energy experiments performed and allow for an important criterion in future rational drug design. Mutation studies of proposed key interaction regions of the drugs may be performed within a free energy of perturbation context to gain valuable insight into drug function. It is important to note that the rapid and reversible binding behaviour of the *Vinca* alkaloids may be crucial to their activity and therefore congeners that both bind too strongly in addition to those that bind too weakly may not be effective.

The *Vinca* alkaloids are considered to act in two ways: 1) the slowing of rates of polymerization and depolymerization of microtubules; and, 2) the curling of microtubule

protofilaments causing rapid depolymerization. The current study addressed the first method of action, but not the second. To determine the geometric effect of bound *Vinca* alkaloids on polymerized microtubules, studies of *Vinca* alkaloids 'sandwiched' between a beta tubulin unit and an alpha tubulin unit could prove fruitful. The 1Z2B structure contains a VLB unit at the interface between two tubulin heterodimers and may be a reasonable starting structure for such 'sandwiched' simulations. However, care would need to be taken in order to not predispose the geometry of the complex: the 1Z2B complex is stabilized by a stathmin-like domain and this appears to induce a bent shape in the complex.

Lastly, in what may prove to be the most interesting and challenging study to be proposed, the human metabolism and metabolites of the *Vinca* alkaloids could be studied using the techniques used here and proposed previously in this section. Several different metabolites of the *Vinca* alkaloids have been identified in a variety of experimental and clinical studies, ranging from basic hydrolysis products to more exotic enzymatically-cleaved structures. Each of these metabolites may also bind to tubulin or other targets. Considering potential binding to tubulin and potential drug-isotype specificity alone, these metabolites may either contribute to or be the primary cause of a *Vinca* alkaloid's anti-mitotic properties or may contribute or cause negative side-effects. The study of observed and rationally predicted metabolites through QM calculations, MD simulations, and docking studies could determine many relevant properties including chemical reactivity, molecular charges, and binding modes, if any. These properties would be highly useful in preparation of an ADMET model to include metabolites. Also, rational design could be carried out to prevent or enhance the drugs' metabolism, depending on the action of the metabolite. Based on the importance of the main dihedral angle torsional energy profile, any metabolites that affect the backbone, especially in the upper portion, of a *Vinca* alkaloid may have entirely unique behaviour. For instance, the reduced activity of Catharinine, where the upper piperidine ring is opened, found in [52], may be due to an altered conformational preference.

Bibliography

[1] Bau, R. and Jin, K. K., *Journal of the Chemical Society-Perkin Transactions 1*, 2000, **13**, 2079–2082.

[2] Gigant, B.; Wang, C. G.; Ravelli, R. B. G.; Roussi, F.; Steinmetz, M. O.; Curmi, P. A.; Sobel, A. and Knossow, M., *Nature*, 2005, **435**, 519–522.

[3] Duffin, J., *Canadian Bulletin of Medical History*, 2000, **17**, 155–192.

[4] Johnson, I. S.; Armstrong, J. G.; Burnett, J. P. and Gorman, M., *Cancer Research*, 1963, **23**(8), 1390–1427.

[5] Chabner, B. A. and Roberts, T. G., *Nature Reviews Cancer*, 2005, **5**(1), 65–72.

[6] Kuehne, M. E.; Bornmann, W. G.; Marko, I.; Qin, Y.; LeBoulluec, K. L.; Frasier, D. A.; Xu, F.; Mulamba, T.; Ensinger, C. L.; Borman, L. S.; Huot, A. E.; Exon, C.; Bizzarro, F. T.; Cheung, J. B. and Bane, S. L., *Organic and Biomolecular Chemistry*, 2003, **1**(12), 2120–2136.

[7] Guéritte, F. and Fahy, J., *Anticancer Agents From Natural Products*, CRC Press, 2005.

[8] Fahy, J., *Current Pharmaceutical Design*, 2001, **7**(13), 1181–1197.

[9] Armstrong, J. G.; Dyke, R. W.; Fouts, P. J.; J., H. J.; Jansen, C. J. and Peabody, A. M., *Cancer Research*, 1967, **27**(Part 1), 221–227.

[10] Pierre, A.; Lavielle, G.; Hautefaye, P.; Seurre, G.; Leonce, S.; Saintdizier, D.; Boutin, J. A. and Cudennec, C. A., *Anticancer Research*, 1990, **10**(1), 139–144.

[11] Hill, B. T., *Current Pharmaceutical Design*, 2001, **7**(13), 1199–1212.

[12] Lavielle, G.; Hautefaye, P.; Schaeffer, C.; Boutin, J. A.; Cudennec, C. A. and Pierre, A., *Journal of Medicinal Chemistry*, 1991, **34**(7), 1998–2003.

[13] Jordan, M. A.; Himes, R. H. and Wilson, L., *Cancer Research*, 1985, **45**(6), 2741–2747.

[14] Sweeney, M. J.; Boder, G. B.; Cullinan, G. J.; Culp, H. W.; Daniels, W. D.; Dyke, R. W.; Gerzon, K.; Mcmahon, R. E.; Nelson, R. L.; Poore, G. A. and Todd, G. C., *Cancer Research*, 1978, **38**(9), 2886–2891.

[15] Langlois, N.; Gueritte, F.; Langlois, Y. and Potier, P., *Journal of the American Chemical Society*, 1976, **98**(22), 7017–7024.

[16] Mangeney, P.; Andriamialisoa, R. Z.; Lallemand, J. Y.; Langlois, N.; Langlois, Y. and Potier, P., *Tetrahedron*, 1979, **35**(18), 2175–2179.

[17] Rai, S. S. and Wolff, J., *Journal of Biological Chemistry*, 1996, **271**(25), 14707–14711.

[18] Moncrief, J. W. and Lipscomb, W. N., *Journal of the American Chemical Society*, 1965, **87**(21), 4963–4964.

[19] Hunter, B. K.; Dhall, L. and Sanders, J. K. M., *Journal of the Chemical Society-Perkin Transactions 1*, 1983, **3**, 657–665.

[20] Gaggelli, E.; Valensin, G.; Stolowich, N. J.; Williams, H. J. and Scott, A. I., *Journal of Natural Products*, 1992, **55**(3), 285–293.

[21] Allinger, N. L., *Journal of the American Chemical Society*, 1977, **99**(25), 8127–8134.

[22] Andrews, C. W.; Wisowaty, J.; Davis, A. O.; Crouch, R. C. and Martin, G. E., *Journal of Heterocyclic Chemistry*, 1995, **32**(3), 1011–1017.

[23] Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T. and Still, W. C., *Journal of Computational Chemistry*, 1990, **11**(4), 440–467.

[24] Allinger, N. L.; Yuh, Y. H. and Lii, J. H., *Journal of the American Chemical Society*, 1989, **111**(23), 8551–8566.

[25] Lii, J. H. and Allinger, N. L., *Journal of the American Chemical Society*, 1989, **111**(23), 8566–8575.

[26] Lii, J. H. and Allinger, N. L., *Journal of the American Chemical Society*, 1989, **111**(23), 8576–8582.

[27] Newton, D. W. and Kluza, R. B., *Drug Intelligence and Clinical Pharmacy*, 1978, **12**(9), 546–554.

[28] Jordan, M. A. and Wilson, L., *Nature Reviews Cancer*, 2004, **4**(4), 253–265.

[29] Huzil, J. T.; Luduena, R. F. and Tuszynski, J., *Nanotechnology*, 2006, **17**(4), S90–S100.

[30] Lowe, J.; Li, H.; Downing, K. H. and Nogales, E., *Journal of Molecular Biology*, 2001, **313**(5), 1045–1057.

[31] Humphrey, W., D. A. and Schulten, K., *Journal of Molecular Graphics*, 1996, **14**(1), 33–38.

[32] Mitchison, T. and Kirschner, M., *Nature*, 1984, **312**, 237–242.

[33] Walker, R. A.; Obrien, E. T.; Pryer, N. K.; Soboeiro, M. F.; Voter, W. A.; Erickson, H. P. and Salmon, E. D., *Journal of Cell Biology*, 1988, **107**(4), 1437–1448.

[34] Wilson, L. and Jordan, M. A., *Chemistry and Biology*, 1995, **2**(9), 569–573.

[35] Toso, R. J.; Jordan, M. A.; Farrell, K. W.; Matsumoto, B. and Wilson, L., *Biochemistry*, 1993, **32**(5), 1285–1293.

[36] Jordan, M. A. and Wilson, L., *Biochemistry*, 1990, **29**(11), 2730–2739.

[37] Chatterjee, S. K.; Laffray, J.; Patel, P.; Ravindra, R.; Qin, Y.; Kuehne, M. E. and Bane, S. L., *Biochemistry*, 2002, **41**(47), 14010–14018.

[38] Mitra, A. and Sept, D., *Biochemistry*, 2004, **43**(44), 13955–13962.

[39] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D. and Kale, L., *Journal of Computational Chemistry*, 2005, **26**, 1781–1802.

[40] Mackerell, A. D.; Feig, M. and Brooks, C. L., *Journal of Computational Chemistry*, 2004, **25**(11), 1400–1415.

[41] MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; D., E. J.; Field, M. J.; Fischer, S.; Gao, ; J., ; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D. and Karplus, M., *Journal of Physical Chemistry B*, 1998, **102**(18), 3586–3616.

[42] Morris, G.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K. and J., O. A., *Journal of Computational Chemistry*, 1998, **19**(14), 1639–1662.

[43] Clark, M.; Cramer, R. D. and Vanopdenbosch, N., *Journal of Computational Chemistry*, 1989, **10**(8), 982–1012.

[44] Gasteiger, J. and Marsili, M., *Tetrahedron*, 1980, **36**(22), 3219–3228.

[45] Gasteiger, J. and Marsili, M., *Tetrahedron Letters*, 1978, **34**, 3181–3184.

[46] Sackett, D. L., *Biochemistry*, 1995, **34**(21), 7010–7019.

[47] Kruczynski, A.; Barret, J. M.; Etievant, C.; Colpaert, F.; Fahy, J. and Hill, B. T., *Biochemical Pharmacology*, 1998, **55**(5), 635–648.

[48] Rowinsky, E. K. and Tolcher, A. W., *Chemotherapy Source Book*, Lippincott Williams and Wilkins, 2001; chapter 23; 3rd ed.

[49] Owellen, R. J.; Hartke, C. A. and Hains, F. O., *Cancer Research*, 1977, **37**(8), 2597–2602.

[50] Zhou, X. J.; Zhoupan, X. R.; Gauthier, T.; Placidi, M.; Maurel, P. and Rahmani, R., *Biochemical Pharmacology*, 1993, **45**(4), 853–861.

[51] Jehl, F.; Quoix, E.; Leveque, D.; Pauli, G.; Breillout, F.; Krikorian, A. and Monteil, H., *Cancer Research*, 1991, **51**(8), 2073–2076.

[52] Elmarakby, S. A.; Duffel, M. W.; Goswami, A.; Sariaslani, F. S. and Rosazza, J. P. N., *Journal of Medicinal Chemistry*, 1989, **32**(3), 674–679.

[53] Magnus, P.; Ladlow, M. and Elliott, J., *Journal of the American Chemical Society*, 1987, **109**(25), 7929–7930.

[54] Fujita, E. and Nagao, Y., *Bioorganic Chemistry*, 1977, **6**(3), 287–309.

[55] Castle, M. C. and Mead, J. A. R., *Biochemical Pharmacology*, 1978, **27**(1), 37–44.

[56] Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Wiley, 2nd ed., 2004.

[57] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W. and Klein, M. L., *Journal of Chemical Physics*, 1983, **79**(2), 926–935.

[58] Bayly, C. I.; Cieplak, P.; Cornell, W. D. and Kollman, P. A., *Journal of Physical Chemistry*, 1993, **97**(40), 10269–10280.

[59] Szabo, A. and Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Collier Macmillan, 1982.

[60] Pople, J. A.; Santry, D. P. and Segal, G. A., *Journal of Chemical Physics*, 1965, **43**(10P2), S129.

[61] Pople, J. A. and Segal, G. A., *Journal of Chemical Physics*, 1965, **43**(10P2), S136–S149.

[62] Mataga, N. and Nishimoto, K., *Zeitschrift für Physikalische Chemie*, 1957, **13**, 140.

[63] Pariser, R. and Parr, R. G., *Journal of Chemical Physics*, 1953, **21**(3), 466–471.

[64] Pariser, R. and Parr, R. G., *Journal of Chemical Physics*, 1953, **21**(5), 767–776.

[65] Huckel, E., *Zeitschrift für Physik*, 1931, **70**(204).

[66] Pariser, R. and Parr, R. G., *Journal of Chemical Physics*, 1953, **21**(3), 466–471.

[67] Pariser, R. and Parr, R. G., *Journal of Chemical Physics*, 1953, **21**(5), 767–776.

[68] Pople, J. A., *Transactions of the Faraday Society*, 1953, **49**(12), 1375–1385.

[69] Pople, J. A. and Segal, G. A., *Journal of Chemical Physics*, 1966, **44**(9), 3289–3296.

[70] Pople, J. A.; Beveridg.Dl, and Dobosh, P. A., *Journal of Chemical Physics*, 1967, **47**(6), 2026–2033.

[71] Bingham, R. C.; Dewar, M. J. S. and Lo, D. H., *Journal of the American Chemical Society*, 1975, **97**(6), 1285–1293.

[72] Dewar, M. J. S. and Thiel, W., *Journal of the American Chemical Society*, 1977, **99**(15), 4899–4907.

[73] Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F. and Stewart, J. J. P., *Journal of the American Chemical Society*, 1985, **107**(13), 3902–3909.

[74] Granovsky, A. A., PC-GAMESS Version 7.0.

[75] Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.and Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S. and Kollman, P. A., AMBER 9, University of California, San Francisco.

[76] Dewar, M. J. S. and Zoebisch, E. G., *Theochem-Journal of Molecular Structure*, 1988, **49**, 1–21.

[77] Stewart, J. J. P., *Journal of Computational Chemistry*, 1989, **10**(2), 221–264.

[78] Stewart, J. J. P., *Journal of Computational Chemistry*, 1989, **10**(2), 209–220.

[79] Rocha, G. B.; Freire, R. O.; Simas, A. M. and Stewart, J. J. P., *Journal of Computational Chemistry*, 2006, **27**(10), 1101–1111.

[80] Repasky, M. P.; Chandrasekhar, J. and Jorgensen, W. L., *Journal of Computational Chemistry*, 2002, **23**(16), 1601–1622.

[81] www.openmopac.net/mopac2007.html.

[82] Personal Communication with Jonathan Mane, PhD Candidate in Prof. Klobukowski's group.

[83] Dewar, M. J. S.; Jie, C. X. and Yu, J. G., *Tetrahedron*, 1993, **49**(23), 5003–5038.

[84] Stewart, J. J. P., *Journal of Computational Chemistry*, 1991, **12**(3), 320–341.

[85] Ferguson, D. M.; Gould, I. R.; Glauser, W. A.; Schroeder, S. and Kollman, P. A., *Journal of Computational Chemistry*, 1992, **13**(4), 525–532.

[86] Mulliken, R. S., *Journal of Chemical Physics*, 1955, **23**(10), 1833–1840.

[87] Momany, F. A., *Journal of Physical Chemistry*, 1978, **82**(5), 592–601.

[88] Cox, S. R. and Williams, D. E., *Journal of Computational Chemistry*, 1981, **2**(3), 304–323.

[89] Wang, J. M.; Cieplak, P. and Kollman, P. A., *Journal of Computational Chemistry*, 2000, **21**(12), 1049–1074.

[90] Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M. and Kollman, P., *Journal of Computational Chemistry*, 2003, **24**(16), 1999–2012.

[91] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W. and Kollman, P. A., *Journal of the American Chemical Society*, 1995, **117**(19), 5179–5197.

[92] Jakalian, A.; Bush, B. L.; Jack, D. B. and Bayly, C. I., *Journal of Computational Chemistry*, 2000, **21**(2), 132–146.

[93] Jakalian, A.; Jack, D. B. and Bayly, C. I., *Journal of Computational Chemistry*, 2002, **23**(16), 1623–1641.

[94] Flyvbjerg, H. and Petersen, H. G., *Journal of Chemical Physics*, 1989, **91**(1), 461–466.

[95] Spartan '04 for Macintosh. Wavefunction Inc., Irvine, CA.

[96] Halgren, T. A., *Journal of Computational Chemistry*, 1996, **17**(5-6), 490–519.

[97] Halgren, T. A., *Journal of Computational Chemistry*, 1996, **17**(5-6), 520–552.

[98] Halgren, T. A., *Journal of Computational Chemistry*, 1996, **17**(5-6), 553–586.

[99] Halgren, T. A. and Nachbar, R. B., *Journal of Computational Chemistry*, 1996, **17**(5-6), 587–615.

[100] Halgren, T. A., *Journal of Computational Chemistry*, 1996, **17**(5-6), 616–641.

[101] Head, J. D. and Zerner, M. C., *Chemical Physics Letters*, 1985, **122**(3), 264–270.

[102] Banerjee, A.; Adams, N.; Simons, J. and Shepard, R., *Journal of Physical Chemistry*, 1985, **89**(1), 52–57.

[103] Nam, K.; Gao, J. L. and York, D. M., *Journal of Chemical Theory and Computation*, 2005, **1**(1), 2–13.

[104] Ryckaert, J. P.; Ciccotti, G. and Berendsen, H. J. C., *Journal of Computational Physics*, 1977, **23**(3), 327–341.

[105] Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A. and Case, D. A., *Journal of Computational Chemistry*, 2004, **25**(9), 1157–1174.

[106] Meagher, K. L.; Redman, L. T. and Carlson, H. A., *Journal of Computational Chemistry*, 2003, **24**(9), 1016–1025.